(12)

FG.

ADA034770

RESEARCH ON APTITUDES: A PROGRESS REPORT

RICHARD E. SNOW

TECHNICAL REPORT NO. 1
APTITUDE RESEARCH PROJECT
SCHOOL OF EDUCATION
STANFORD UNIVERSITY

Sponsored by

Personnel and Training Research Programs
Psychological Sciences Division
Office of Naval Research

and

Advanced Research Projects Agency

under

Contract No. N00014-75-C-0882

SEPTEMBER 1976

S/C
391909

RESEARCH ON APTITUDES: A PROGRESS REPORT

RICHARD E. SNOW

TECHNICAL REPORT #1
APTITUDE RESEARCH PROJECT
SCHOOL OF EDUCATION
STANFORD UNIVERSITY

Sponsored by
Personnel and Training Research Programs
Psychological Sciences Division
Office of Naval Research
and
Advanced Research Projects Agency
under
Contract No. N00014-75-C-0882

SEPTEMBER 1976

The preparation of this report was supported in part by a grant from the Spencer Foundation, in addition to Contract No. N00014-75-C-0882, Office of Naval Research. It will appear as a chapter in L. Shulman (Ed.), Review of Research in Education, 1977, Vol. 4, in press. One portion of the report was also included in presentations to the American Educational Research Association, April 1974, and to the First Annual Delaware Symposium on Curriculum, Instruction, and Learning, May 1974.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER 1	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER 9
4. TITLE (and Subtitle) 6 Research on aptitudes: A progress report.		5. TYPE OF REPORT & PERIOD COVERED Technical Report, 1 15
		6. PERFORMING ORG. REPORT NUMBER 1
7. AUTHOR(s) 10 Richard E. Snow		8. CONTRACT OR GRANT NUMBER(s) 15 N00014-75-C-0882 NEW
9. PERFORMING ORGANIZATION NAME AND ADDRESS School of Education Stanford University Stanford, California 12 68p.		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS NR 154-376
11. CONTROLLING OFFICE NAME AND ADDRESS Personnel and Training Research Programs Psychological Sciences Dividion, ONR, 458		12. REPORT DATE 11 September 1976
		13. NUMBER OF PAGES 60
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) 14 TR-1		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

UNLIMITED

DISTRIBUTION STATEMENT A
Approved for public release;
Distribution Unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

UNLIMITED

18. SUPPLEMENTARY NOTES

This research jointly sponsored by the Office of Naval Research and the Defense Advanced Research Projects Agency.
Also to appear as a chapter in Shulman, L. (Ed.) Review of Research in Education, Vol. 4.

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Aptitude, instructional methods, aptitude-treatment interaction, individual differences in ability, personality, and learning

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

A review of current literature on the relation of aptitudes to learning outcomes under different instructional methods or conditions. Included are: aptitude-instructional treatment interaction studies, notes on new methodological problems, and studies of individual differences in laboratory learning arrangements. Aptitudes emphasized include abilities (general, fluid-analytic, crystallized) and personality traits (anxiety, achievement via independence and achievement via conformity). Complex

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-014-6601

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

391909

1B

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)

interactions among ability, personality, and instructional treatment variables are demonstrated. Implications for further research are discussed.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)

Individual differences among learners constitute an important class of variables for research on instruction. Their study has been of interest at least since Binet, because measures of these variables, often called "aptitudes", usually predict learning outcome. There is renewed interest in this fact today because aptitudes now often appear to interact with instructional conditions, i.e., to relate differently to learning outcome under different instructional treatments. Practical interest stems from the possibility that such interactions can be used to adapt instruction to fit different learners optimally, since previous attempts at individualizing instruction have generally failed to eliminate individual differences in learning outcome. Such interactions are of theoretical interest because they demonstrate construct validity for aptitude and learning measures in a new way, and raise the possibility that common processes underlie both kinds of variables. They suggest that neither aptitude constructs nor learning processes can be fully understood without reference to the other.

Aptitude-treatment interactions (ATI) have been the subject of many studies in recent years, and a wide variety of ATI findings are now on hand. From their review of this research, Cronbach and Snow (in press) concluded that the existence of ATI as phenomena has been clearly established. But, while some ATI findings are plausible and some are replicable, few are well understood and none are yet applicable to instructional practice.

The volume of ATI studies can be arrayed along a continuum from laboratory experiments on individual differences in learning, through small-scale and middle-range instructional experiments, to large curriculum evaluations, naturalistic comparisons, and empirical case studies. If one constructed a frequency distribution of such studies along this continuum, it would probably appear more or less normal in form; most studies would fall in the middle range. And, the same form of distribution might be expected for instructional experiments in general, not just for ATI studies. Thus, the instructional psychology popular today consists of short-term experiments with a few controlled instructional variables aimed at testing fairly simple propositions. These studies seek compromise between the need for instructional relevance and the need for experimental control. But most attain neither the descriptive value of large-scale, long-term naturalistic research nor the precision and process analysis of the laboratory; inconsistencies abound in their results. Cronbach and Snow (in press) concluded that the

middle-range studies were leading neither to theory nor to generalizations useful in practice. They recommended that future research pay increased attention to:

1) The examination of the most plausible ATI hypotheses in large scale, long-duration, real-school studies. This would allow a consolidation of efforts to establish a few ATI hypotheses in settings where they might actually be used. The emphasis in the design of such research would be on representativeness (Snow, 1974) and description (Cronbach, 1975), rather than on laboratory-like control.

2) The development of methodology capable of handling the complexities of such research. This effort would deemphasize the familiar significance testing habits of researchers in favor of the description and analysis of complex relationships (Cronbach & Snow, in press; Cronbach & Webb, 1975; Cronbach, 1976).

3) The development of a laboratory science for the analysis of aptitude tests and learning tasks, and the ATI constructs based on them (Snow, 1976). This would complement the instructional studies with process analyses to provide ideas about possible underlying mechanisms. Embodied in newly understood and/or newly designed aptitude measures, these ideas might then be conveyed to research in the real instructional settings where probable, practically useful ATI can be examined and used.

The present review must be highly selective with respect to each of these three lines of continuing research. It will avoid repeating material available in the above cited sources wherever possible. Since the previous review in this series by Berliner and Cahen (1973) and the Cronbach-Snow book, many new ATI studies have appeared. No attempt will be made here to catalog this heterogeneous collection; only studies bearing on a few major ATI hypotheses and methodological developments are reviewed. Suggestions for further research are made along the way.

Definitional Matters

Before proceeding, it may be helpful here to deal with several definitional concerns that have confused thinking and writing in the ATI field. Some of these are touched upon in Cronbach and Snow (in press); some are not.

The first concern is with the term "aptitude" itself. Cronbach and Snow (in press) defined aptitude as "...any characteristic of a person that forecasts his probability of success under a given treatment" (p. 6). This purposely broad and pragmatic definition was meant to promote a new kind of instructional research. Personality variables, biographical and other non-test measures, and new kinds of aptitude constructs might predict response to instruction in a given setting, singly or in combination, and hence might be thought of as sources of aptitude for success in that setting. The traditional conception of aptitude for school learning, as represented exclusively by "scholastic aptitude tests" or "general ability tests" was unnecessarily limiting. Other work in differential psychology has developed more specialized multivariate conceptions of aptitude for success in specialized activities, and these span the artificial distinctions between achievement, ability, personality, etc.: mechanical knowledge is one aptitude for technical work, along with various psychomotor skills; some esthetic sensitivities, along with drawing and perceptual skills, are aptitudes in art; a certain degree of compulsivity may be an additional aptitude for computer programming; even height is an aptitude for basketball. There is no a priori reason why aptitude for success in instructional activities should be less complex, diversified, or specialized than it is in these other pursuits.

Other writers, hoping to avoid potential confusions, have preferred terms like "attribute-treatment interaction" or "trait-treatment interaction". But these are hollow terms--empty of substantive meaning. "Aptitude", on the other hand, is a substantive concept in educational psychology. We should expect continued research to alter and elaborate the meaning of such concepts, just as it alters and elaborates concepts of instructional method. Further, individual aptitude might be expected to develop or change with continued experience in a given kind of instruction; "trait" and "attribute" imply permanence. Most importantly, we can hope there will someday be a theory of "aptitude"; there can never be a theory of "trait" or of "attribute". The present writer thus persists in advocating use of the substantive term. Research on aptitude for learning is the study of individual differences in learning and learning-related processes, particularly as these vary and covary under different instructional conditions.

A second issue arose over the distinction between intrinsic and extrinsic individual differences relevant to learning. The former take their definitions directly from measures of learning processes; the latter do not.

The extrinsic category thus includes most aptitude constructs available from differential psychological research and there is the suggestion that these are not process-based at all. The distinction was introduced by Jensen (1967) simply to classify prior studies, and was later discarded by him. Although some researchers have maintained the contrast as substantively important, it is of doubtful merit. First, to assert that aptitude measures are not concerned with psychological processes is absurd, even though differential psychology has not yet been much concerned with the development of process theories. When an aptitude measure relates differently to learning measures under different learning conditions, there is the implication that that aptitude is fundamentally involved in learning processes under one or both of those conditions. Whether the processes represented by the aptitude measure are related to, or are the same as, those represented by the learning measure to which it relates is an empirical and theoretical matter. There is considerable evidence, for example, that general mental tests represent the ability to learn in conventional instruction, and some further evidence that personality constructs combine with mental ability to magnify or dampen this relation. The task for further research is to build a more detailed process theory of such relations. It helps such research not at all to prejudge the matter with arbitrary classifications.

Relevant here also is the optimistic view that process-oriented research on individual differences in learning will provide "new aptitudes" (Glaser, 1972), different in kind from the "old aptitudes". This is an important possibility. Since individual differences in mental performances almost always correlate however (Guttman, 1976), it is more likely that the new and the old will differ in form more than in kind, and that an improved, integrated conception of human cognition will need to be built on their combination. Both kinds of constructs always must be included in such research anyway, since new constructs cannot be defined without demonstrating discriminant validity with respect to existing constructs (Campbell and Fiske, 1959). This admonition applies as much to the old artificial distinctions in differential psychology between ability, style, and personality, as it does to new artificial distinctions between what is measured by tests and what is measured by laboratory task parameters. We can strive for, and expect, new aptitude constructs, but these are likely to be woven in large part from the threads of existing cloth .

A third matter can be dispensed with quickly. Early discussions of ATI,

including those of the present writer, emphasized only disordinal interactions (where regression lines intersect within the aptitude range). Instances of ordinal interactions were ignored, or were classed with instances of no interaction. But ordinal and disordinal interactions can have similar practical implications, depending on other aspects of instruction, e.g., costs. More importantly, both have the same theoretical implications. The distinction is thus unimportant for the purposes of future research.

Finally, the ATI approach has been defined by some as relevant only to one narrow form of individualized instruction, i.e., to situations where students can be assigned to alternative instructional treatments. It is now clear, however, that all attempts at individualizing instruction rest explicitly or implicitly on hypothesized ATI. Anytime an instructional prescription for one student differs from that provided another, there is the suggestion that each is best helped toward some common goal by following his or her own prescription rather than someone else's. Further, the evaluation of instructional prescriptions, whether individualized or not, requires an ATI approach even where there is no intent ultimately to assign students to alternative instructional treatments. *In describing any kind of instructional* effect, one must always be able to say whether the description given holds for each student involved. Research on aptitude thus takes a place in more general efforts to build instructional theory. A theory of aptitude is required in the second, or "description of initial state", part of Glaser's (1976) four-part conception of prescriptive theory for instruction, and the methodology of ATI research fills out the fourth part--"assessment of instructional effects".

Instructional Studies

The present review of instructional studies concentrates on two hypotheses. One of these asserts that individual differences in anxiety (A_x), achievement via independence (A_i), and achievement via conformity (A_c) interact with instructional treatments differing in the degree of structure and participation provided the student. This will be referred to simply as the $A_iA_cA_x$ complex. The other concerns general mental ability (G) and the extent to which its relation to learning outcome varies with the information processing burden placed on the learner by the instructional treatment. Since general ability tests typically

combine types of items that have been theoretically distinguished (by Cattell, 1971, and Horn, 1976) as representing crystallized ability (G_c), fluid ability (G_f) and spatial visualization ability (G_v), this hypothesis will be referred to as the $G_cG_fG_v$ complex. The two aptitude complexes are probably not orthogonal.

The $A_iA_cA_x$ complex. A large number of ATI studies that used personality constructs as aptitudes were summarized by Cronbach and Snow (in press) under the general rubric of "constructive" vs. "defensive" motivation. It was hypothesized that constructively motivated students, those ready to take confident, self-directing initiative in learning, would profit in less directive instructional situations that allow and encourage student initiative; defensively motivated students, those more anxious, dependent, or conforming, would require situations tailored to provide a more supportive external structure, with less demand for independent action, within which such students could work effectively. This was admittedly a gross and oversimplified contrast, but several strong studies had provided its core support. These certainly were worth more generalized attention.

Several of the core studies were contributed by Domino. In one (Domino, 1968), college students, matched on sex and nonverbal ability, were classified as High-High, Low-Low, Low-High, and High-Low on the basis of their scores on the A_i and A_c scales from the California Psychological Inventory. Domino then interviewed instructors of every course taken by the students, to classify the courses as "encouraging conformity" or "encouraging independence". Variables such as emphasis on objectives, memorization, attendance, etc., vs. independent reading, student discussion, informal evaluation, etc., were used to direct the classification. It was found that students showing a High A_i- Low A_c, profile achieved better grades in courses where they were encouraged to be independent. Those with a Low A_i - High A_c profile obtained higher average grades in the more structured courses. High - High students did better than Low - Low students in general.

Domino (1971) then followed with a formal experiment, in which High A_i- Low A_c and Low A_i - High A_c extreme groups were defined. Here, a single instructor taught four sections of Introductory Psychology, two in a structured style emphasizing conformity and two in a style emphasizing independence. Extreme groups were divided randomly among the two treatments, and several achievement outcome measures were used. Again, students high in A_i and low in A_c did

best with instruction favoring independence; students low in A_i and high in A_c were better off with instruction requiring conformity.

In another line of work, Dowaliby and Schumer (1973) contrasted teacher-centered vs. student-centered teaching in a junior college psychology course. In the former treatment the teacher carried the communication burden, with little participation by students. In the latter treatment, students were encouraged to ask questions and contribute ideas. The aptitude of interest was the Taylor Manifest Anxiety Scale (A_x). There were two examinations as outcome measures. ATI was marked; the more anxious students were better treated by teacher-dominated instruction, while less anxious students achieved more with participative instruction.

Domino (1974) also replicated the Dowaliby-Schumer results. College students in an English course were assigned to teacher-centered or student-centered instruction. A_x again interacted with achievement outcome, showing that outcome was superior for high A_x students in the teacher-centered approach and for low A_x students in the student-centered approach.

These studies along with related research reviewed by Cronbach and Snow (notably a series of studies by McKeachie), are sufficient to establish the importance of ATI in this domain, but the conclusion sustained by their combined results is not really satisfactory. Several questions need answers. Are treatment contrasts characterized as structured vs. participative, or conforming vs. independent, or teacher-centered vs. student-centered similar? In what ways? Is this a college-level phenomenon only, or are similar interactions observable at lower levels of education? Are A_i, A_c, and A_x different faces of the same general motivational construct? How are any or all of these aptitude variables related to G, which has been found in other research to combine with A_x in higher-order interactions? Two more recent investigations have sought to push our understanding of this complex of questions further.

A study by Peterson (1976) obtained striking interactions which correspond roughly to Domino's results but complicate the Dowaliby-Schumer interpretation appreciably. They also show the constructive-defensive hypothesis of Cronbach-Snow to have been overly simple. Peterson defined four treatment conditions to distinguish between the Domino emphasis on teacher structure and the Dowaliby-Schumer emphasis on student participation. One teacher taught a two-week unit on alienation to each of four high school classes in Social Studies (N=94), assigned at random to a 2x2 design. The treatments were defined as high structure and high participation (HS:HP), high structure and low participation (HS:LP), low structure and high participation (LS:HP), and low

structure and low participation (LS:LP). Peterson trained the teacher in these four styles using videotape models. A description of teacher behavior in the four treatments was built into the models, and later verified by classroom observations. Thus a fairly detailed account of teacher behavior in each treatment was available. She also included four aptitude constructs: G (based on a Verbal Comprehension Test), A_x (combining the Children's Manifest Anxiety Scale and the Spielberger Measure), A_c plus A_i, and A_c minus A_i. The latter two variables were defined to distinguish general motivation toward achievement from special orientation toward independence vs. conformity. (It should be noted here that aptitudes entered the regression equation in the order G, A_x, A_c plus A_i, A_c minus A_i, so the latter are partial variates.) Outcome measures were immediate and delayed multiple choice tests, an essay test, and several attitude scales.

Multiple regression analysis on immediate achievement showed a main effect for G, no main effect for treatment, and two substantial ATI effects. One of these, for A_c minus A_i, replicated the Domino findings, as shown in Figure 1. The LS:LP treatment was best for students high in A_i and low in A_c. Next best for these students was the HS:HP treatment. For students with the opposite pattern, i.e., high A_c and low A_i, the HS:LP treatment was superior. Next best, at least for those in this group with high general motivation, was LS:HP. Peterson concluded that conforming students needed one clear strategy to which they could conform. This would preferably be teacher structure. Without that, the structure provided by peer participation might substitute, at least for those high enough also on A_i. Having structure provided by neither teacher nor students was particularly inadequate for these learners. The independent students thrived in this latter environment, particularly if they were quite low on A_c. Presumably, they preferred and were able to provide their own structure. For High A_i students also showing some degree of A_c, however, HS:HP and LS:HP were not inadequate treatments. Results for the other cognitive outcome measures were similar, although a shift among independent and highly motivated students was noted on the essay test; the high participation conditions were most beneficial here. Also, ATI effects faded somewhat at retention.

Figure 1 here

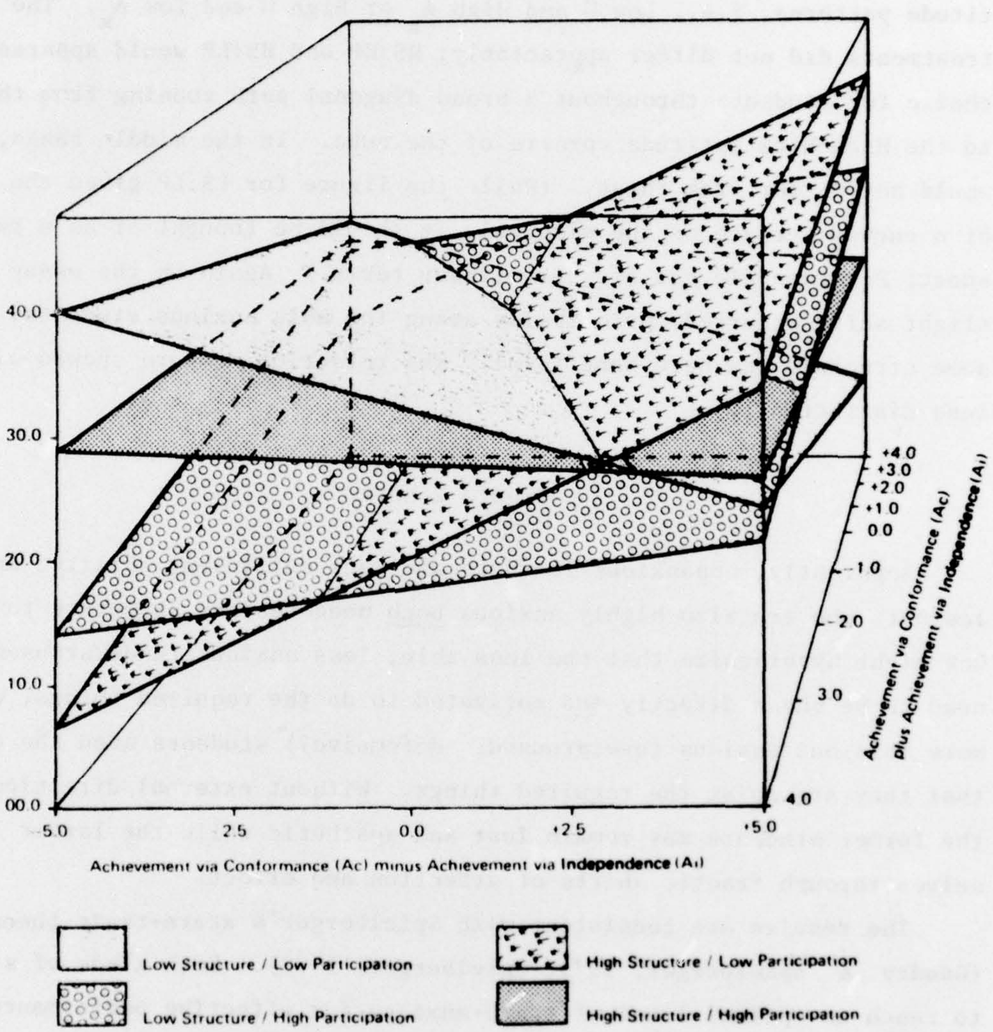


Figure 1. Regression planes for four treatments using A_c plus A_i and A_c minus A_i as aptitudes and multiple-choice achievement as outcome (after Peterson, 1976).

The other ATI involved G and A_x. G alone gave no ATI, nor did A_x alone. But the multiplicative combination of G and A_x yielded the result shown in Figure 2 (the four planes are separated here to simplify what would be a highly complex figure). The operative treatment in the interaction is LS:LP. It was the worst treatment by far for students low in both aptitudes or high in both. It was the best treatment by far for students with the opposite aptitude patterns, i.e., Low G and High A_x or High G and Low A_x. The other treatments did not differ appreciably; HS:HP and HS:LP would apparently be the choice for students throughout a broad diagonal path running from the Low-Low to the High-High aptitude corners of the cube. In the middle range, LS:LP would not differ from these. (While the figure for LS:LP gives the impression of a curved surface on the diagonal, it should be thought of as a twisted sheet; Peterson did not fit curvilinear terms.) Again on the essay test a slight shift occurred, principally among the able anxious students, for whom some structure was here beneficial. The retention measure showed similar but less distinct ATI.

Figure 2 here

Apparently, nonanxious learners who lack sufficient ability, and able learners who are also highly anxious both need teacher structure to do well. One might hypothesize that the less able, less anxious (underaroused?) students need to be shown directly and motivated to do the required things, while the more able but anxious (overaroused? defensive?) students need the certainty that they are doing the required things. Without external direction on this, the former students may remain lost and apathetic while the latter lose themselves through frantic shifts of attention and effort.

The results are consistent with Spielberger's state-trait theory of anxiety (Gaudry & Spielberger, 1971; Spielberger, 1972). Both kinds of students need to reach an optimal level of state-anxiety for effective performance. Teacher direction brings the High G-High A_x students down to this level; teacher demand brings the Low G-Low A_x students up to this level. Able, nonanxious students, on the other hand, are benefitted by absence of formal structure and participation, perhaps for the same reasons that independent students are benefitted by this treatment. In a moderately difficult task, these students are aroused enough and able enough to organize their own work effectively. It is less

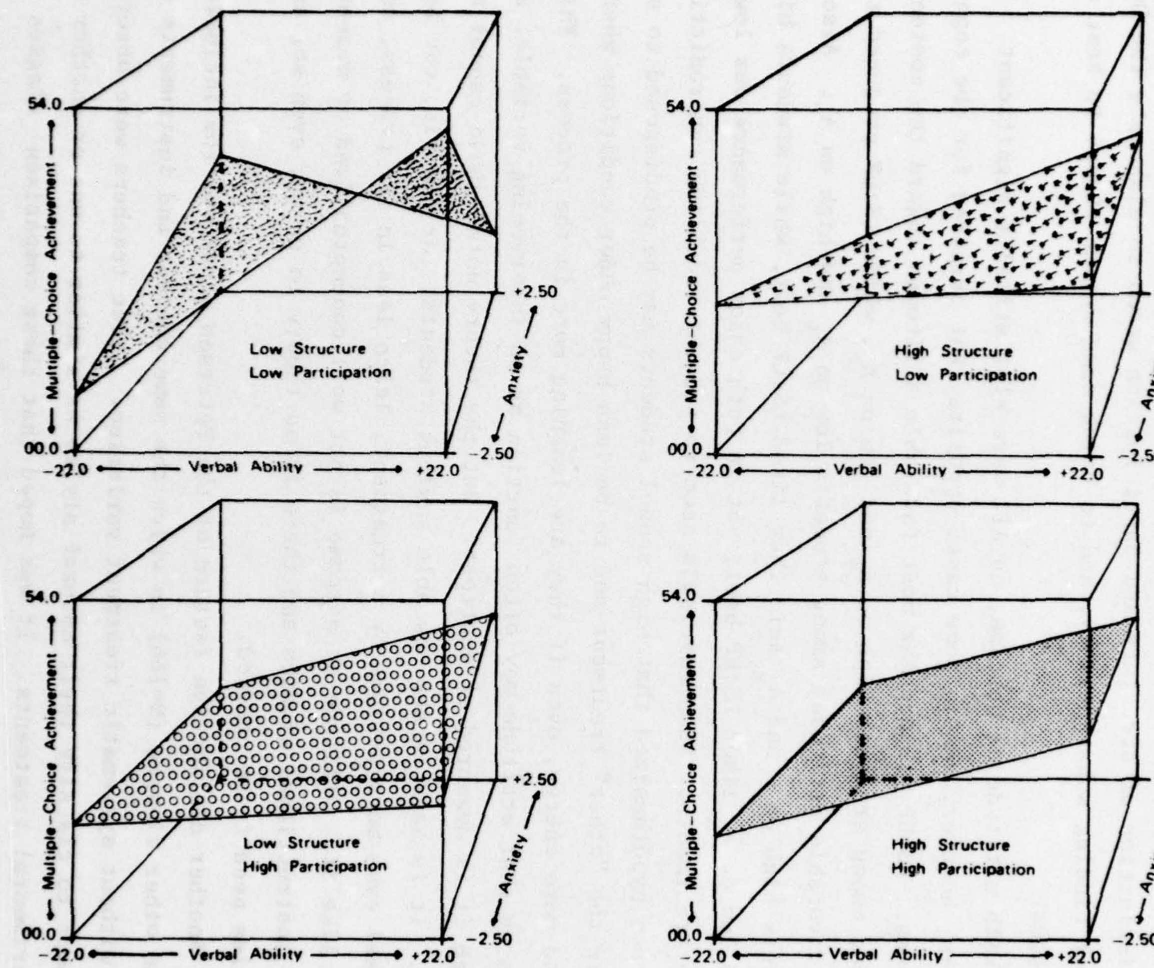


Figure 2. Regression planes for four treatments using G and A_x as aptitudes and multiple-choice achievement as outcome (after Peterson, 1976).

clear why LS:LP serves the less able, anxious students well. Following the Speilberger theory, one could say that anxiety present in these students already serves an optimal activating function so situational demand is unnecessary, even debilitating. Without forced structure or participation, these students can work inconspicuously. In short, the inclusion of ability and the distinction of structure and participation seems to elaborate the Dowaliby-Schumer finding substantially, and to bring some extant theory to bear on ATI processes.

With attitude as outcome, no ATI were statistically significant. Their pattern, however, was in some cases opposite that obtained for the cognitive outcomes. HS:HP yielded the most favorable attitudes toward the content learned among students high on A_i and low on A_c, while LS:LP produced the most favorable attitudes among students low on A_i and high on A_c. Also, students high on G and A_x said they liked LS:LP best, while students high on G and low on A_x liked LS:HP best; yet in both cases performance was lowest in these treatments. Such reversals seem to present troubling contradictions. Peterson hypothesized that high school students may be predisposed to say they prefer the "other" treatment and to be less happy under conditions where they expend more energy, even if they are learning more in the process. This reminds us that attitude may often function as an intervening variable, although it is measured at posttest. But the theory noted above cannot handle this. It is hard to see how able anxious students, for example, can be made to feel even more anxious by a treatment, learn less in the process, and say they like it. Attitude as outcome is not well conceptualized or measured in many instructional studies and there is no theory to go on; even so, such puzzles need to be checked.

Another distinctive feature of the Peterson study was the inclusion of seven other classes (N=166) in which the same content and instruments were used, but without systematic treatment variation. These teachers were observed in an attempt to classify their natural styles as similar to one or another of the experimental treatments. It was hoped that these comparison classes would serve to test the representativeness of the experimental classes and to crossvalidate their regression equations. The results were somewhat mixed. Some comparison classes gave results consistent with predictions, some did not.

A multidimensional scaling of distance measures among the 11 class regression equations suggested that HS:HP and LS:HP were more similar to other natural classes in the school than were HS:LP and LS:LP. In addition, it was clear that other treatment dimensions were operating in the comparison classes to influence similarity of regression equations obtained in each.

Porteus (1976) was able to study some of the same variables in a year-long investigation covering two subject matters, Economics and Educational Philosophy. Students (N=56) in a private high school were assigned to one of two sections of each course, all taught by the same teacher. In one section, a teacher-centered treatment was used, with required attendance, daily reading assignments and homework, frequent quizzes, and term projects all chosen by the teacher. In-class work was teacher-directed, with previews and summaries. The other section was student-centered; students initiated class discussion, requested quizzes and homework as they wished, chose projects, and reading was assigned in 2-4 week blocks. Outcomes were three achievement tests spaced across each two-term course, an essay measure added to the third of these, and student attitude and course perception questionnaires after each test. Classroom audiotapes, a teacher log, and student perceptions helped describe instruction over the year.

The aptitude battery initially included the same A_i and A_c measures used previously plus the Flexibility Scale from the same inventory, A_x (Spielberger's scale), several measures of G_f or G_c (Necessary Arithmetic Operations, Hidden Figures, SAT-V, and SAT-Q), a questionnaire designed by Myers (1964) as a substitute for projective measures of achievement motivation, and a paragraph completion test used by D. Hunt (1975) to represent a cognitive style construct called "conceptual level". The correlation pattern and small sample size, however, suggested that these variables should be reduced to a smaller number of factors if possible. Porteus chose four orthogonal aptitude factors for the ATI analysis. These were: G (since a distinction between G_f and G_c could not be obtained); A_i (a factor combining A_i and Flexibility); A_c (a factor combining A_c, negative Flexibility, and the achievement motivation questionnaire); and A_x. The expansion of the A_i and A_c constructs by the additional components does not seem to change their meaning appreciably, so labels consistent with prior work have been used here. (Porteus used somewhat different labels.) Also, the conceptual level measure was split by the factor analysis, with

positive loading on G, negative on A_x. This makes sense; the scoring of paragraph completion favors conceptually complex and unrestricted verbal production, as opposed to simple, concrete, restricted production. Ability should enhance, and anxiety inhibit, such production. This hint that one style construct may represent a combination of ability and personality constructs, contrasting High G-Low A_x with Low G-High A_x, should spur further research in this direction.

In Economics, multiple regression analysis showed main effects for G and A_c on all but one outcome. The essay test gave main effects associated with all four aptitudes, with that for A_i the largest. There were no treatment main effects, but several ATI were pronounced. In Education, no effects of any kind were nominally significant, although G gave relatively large main effects and ATI appeared noteworthy. R^2 for the full regression model in Education accounted for only 36% to 51% of the variance; in Economics R^2 ranged from 65% to 88%. This suggests a difference due to subject matter.

For A_i and A_c in Economics, ATI was substantial only on Test 2; High A_i-Low A_c students did best in student-centered instruction, while Low A_i-High A_c students were better off in the teacher-centered condition. This seems consistent with Domino's findings and with Peterson's, if one looks only at her low participation groups. Test 2 was given at the end of the first term, so the timing also corresponds to Domino. But results for the other outcomes did not match earlier findings, and it was the regression plane for the student-centered group that seemed to shift. On Test 1, this treatment was poorest for Low A_c students, with no differences for other students. On Test 3 and the essay, teacher-centered instruction was best for Low A_c students, and High A_c students did best with student-centered instruction. A_i had only slight moderating effects on these trends. In Education, all the results were faint. Test 1 data seemed to match those for Economics Test 2, again replicating Domino and Peterson. But Education Test 2 showed results like those of the third stage measures in Economics, and ATI were absent at the third stage in Education.

Considering ATI for G and A_x, Porteus's results were remarkably similar to Peterson's, in both Economics and Education, if one is willing to equate the student-centered and LS:LP conditions on Test 1 and then make some allowance for student and teacher adaptations across Porteus's very long time scale.

Figures 3 and 4 give these results, for Economics and Education respectively. In each figure, panels a, b, c, and d correspond to Tests 1, 2, 3, and the essay, respectively. The most striking ATI appeared on Economics Test 1. The student-centered treatment assumed the same twisted sheet pattern seen earlier in Figure 2 for LS:LP, while the teacher-centered treatment yielded a sloped regression plane quite similar to the planes for Peterson's other three treatments in Figure 2. As before, Low G-Low A_x and High G-High A_x students ~~did best~~ with teacher-centering; Low G-High A_x and High G-Low A_x students profited more under student-centering. Porteus also did not test for curvilinearity.

Figures 3abcd and 4abcd here

On later Economics outcomes, the shapes of the two regression planes were retained, although the twisted appearance of the student-centered treatment faded somewhat over time. Some changes in elevation of the planes over time also would change the resulting decision rules for assigning students to treatments; this is unimportant here since no such conclusions would be drawn from this study alone anyway, given the extremely small sample size. The similar regression patterns nonetheless give encouragement to the general hypothesis when they match those of other studies.

In Education, there was no ATI at Test 1 (Figure 4a). After that, the same pattern as seen in the Economics data emerged across time, becoming most pronounced on the essay test. Here again, then, High G-Low A_x and Low G-High A_x students did best with student-centering; students low on both aptitudes or high on both did best with teacher-centering.

Analyses of attitude toward the instruction yielded only one noteworthy ATI. Low G-High A_x students in Economics felt better about the instructional method in the student-centered section than their counterparts did in the teacher-centered section. A few contradictory trends like those Peterson found were seen, but there were also instances in which attitude and achievement outcomes showed similar patterns.

The interpretation applied to Peterson's results using Spielberger's theory can be applied here also to account for the general ATI pattern. But Porteus's detailed description must be relied upon beyond this. Evidence from classroom audiotapes, teacher logs, and student questionnaires suggested several factors operating between courses, and within courses over the year, that might

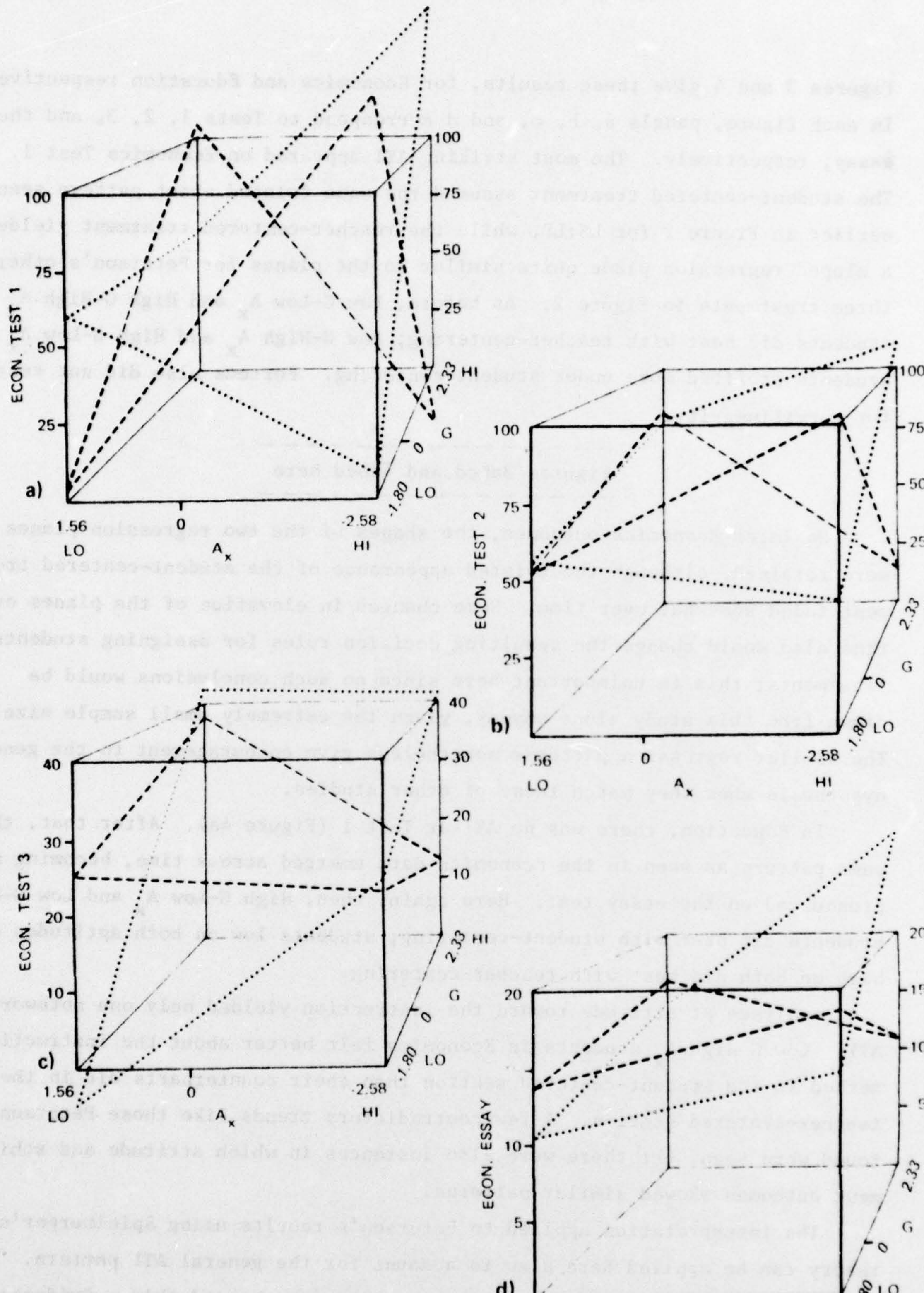


Figure 3. Regression planes for two treatments using G and A_x as aptitudes and Economics Test 1 (a), Test 2 (b), Test 3 (c), and Essay (d) as outcomes (after Porteus, 1976).

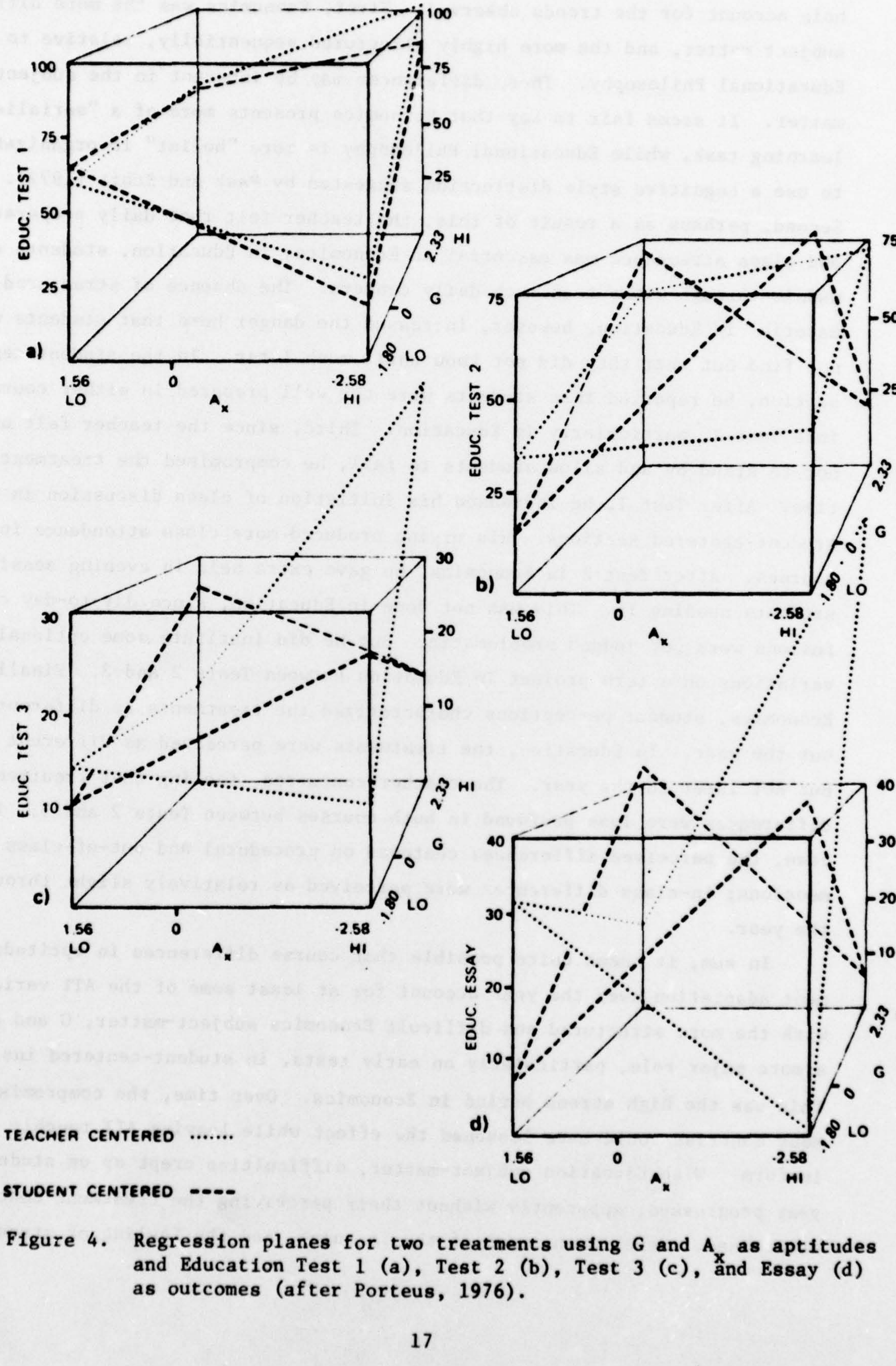


Figure 4. Regression planes for two treatments using G and A_x as aptitudes and Education Test 1 (a), Test 2 (b), Test 3 (c), and Essay (d) as outcomes (after Porteus, 1976).

help account for the trends observed. First, Economics was the more difficult subject matter, and the more highly structured sequentially, relative to Educational Philosophy. These differences may be inherent in the subject matter. It seems fair to say that Economics presents more of a "serialist" learning task, while Educational Philosophy is more "holist" in organization, to use a cognitive style distinction suggested by Pask and Scott (1972). Second, perhaps as a result of this, the teacher felt that daily preparation and class attendance was essential in Economics; in Education, students could muddle through without as much daily concern. The absence of structured text material in Education, however, increased the danger here that students would not find out what they did not know until much later. In the student-centered section, he reported that students were not well prepared in either course before Test 1, particularly in Education. Third, since the teacher felt obligated not to stand by and allow students to fail, he compromised the treatments over time. After Test 1, he increased his initiation of class discussion in the student-centered sections. His urging produced more class attendance in both courses. After Test 2 in Economics, he gave extra help in evening sessions to students needing it. This was not done in Education, since day-to-day confusions were not judged problematic. But he did institute some optional variations on a term project in Education between Tests 2 and 3. Finally, in Economics, student perceptions characterized the treatments as different throughout the year. In Education, the treatments were perceived as different earlier but not later in the year. The teacher concurred, feeling that treatment differences were less profound in both courses between Tests 2 and 3. In either case, the perceived differences centered on procedural and out-of-class dimensions; in-class differences were perceived as relatively slight throughout the year.

In sum, it seems quite possible that course differences in aptitude-treatment adaptation over the year account for at least some of the ATI variations. With the more structured and difficult Economics subject-matter, G and A_x played a more major role, particularly on early tests, in student-centered instruction. This was the high stress period in Economics. Over time, the compromised treatment contrast could have lessened the effect while leaving ATI roughly similar in form. With Education subject-matter, difficulties crept up on students as the year progressed, apparently without their perceiving the treatment contrast being maintained. Here, then, much of the learning, and the feeling of stress, probably

occurred late in the course. It is not surprising then that ATI for G and A_x showed most strongly on the final essay. (See Cronbach and Snow for other studies showing A_x effects varying with stress.) This may account, too, for the confirmation of Domino's A_i A_c hypothesis only on Education Test 1; the High A_i student begins independent learning without prompting. This is inconsistent with the fact that in Economics Domino's effect appeared only on Test 2, unless one says that it takes longer for High A_i to become effective in this difficult course. The teacher did report that his Economics section finally became "truly" student-centered as
report that his Economics section finally became "truly" student-centered as
Test 2 approached. And evening tutoring (for Low A_i students?) was instituted after that. Admittedly, this is all impressionistic description, but it is credible, and it clues the next investigator about variations to expect.

Both the Peterson and Porteus studies were forced to use small samples of students, the Porteus study extremely so within subject-matters, so their statistical power was weak and the regression surfaces may be poorly estimated. But both were strong in the sense of ecological sampling. Peterson contrasted four sharply defined treatments within a school context where seven other naturally varying classes could be studied comparatively. Porteus was able to trace the course of aptitude-treatment adaptation across a year's instruction in two subject-matters. Most importantly, both yielded ATI consistent enough with previous research, and with each other, to verify that the $A_iA_cA_x$ complex is an important source for ATI research.

Yet, understanding of this phenomenon is still not sufficient for theory or practice. Despite obtaining fairly detailed descriptions of treatments, the two studies cannot really map their treatments onto one another. What is needed is an observation system designed specifically to study this ATI complex both within and between school situations. For the present, other impediments to integration remain. The curvilinear functioning of A_x mentioned above needs more detailed checking, the two studies do not define A_i and A_c identically, and neither study gives us a clear picture of how these three constructs fit together. Peterson and Porteus are working now to coordinate their analyses and results on these points. Some further information can be given here on the last matter.

In both studies, there was evidence of positive correlation between G and A_i, and negative correlation between A_c and A_x. Table 1 shows these and other results. (Since Peterson used simple variables rather than factors, the table gives simple results rather than factor combinations for Porteus.

Also Peterson used a verbal ability test alone, so Porteus's SAT-V results are shown rather than those based on her general factor. Both are labeled G_c.) Note that the correlation pattern was similar across studies except for A_c, which seemed to relate to G_c and A_i in the Peterson data but not in the Porteus data. These relations suggest that attempts at tracing out a more detailed and coordinated network among these constructs might pivot on G. It appears to relate to A_i and A_c more highly than these measures do to one another. A_x seems to run counter, particularly to A_c. And, there is much evidence, in this chapter and elsewhere, showing interaction of G and A_x. Further research should capitalize on such relations, seeking convergent and discriminant process descriptions of these dimensions, perhaps using observed classroom behavior and study skills. The reliability estimates given in the main diagonals of Table 1 add the suggestion that A_i and A_c are not well measured by the CPI scales alone. Another useful project for further research would be to build alternative measures for these constructs.

- - - - - - - - - - -

Table 1 about here

- - - - - - - - - - -

Some hypotheses to guide process analysis come from the bottom half of Table 1. Selected correlations are reported from a study by Rutkowski and Domino (1975) relating the CPI scales to an inventory of study skills and habits among college students. (Here, the correlations for A_x are approximated, since CPI contains no A_x scale. The approximation was based on correlations from other CPI scales that form a factor commonly thought to represent the opposite of A_x.) A_c relates rather strongly to all four of these dimensions of study behavior. A_i seems associated with all but delay avoidance. G_c and A_x are apparently concerned only with work methods. Unraveling such relationships would be a useful next step.

<u>The $G_cG_fG_v$ complex.</u> All of the above argues for the addition of G to the $A_iA_cA_x$ string. But G is itself a complex. The hierarchical model of ability organization, as interpreted by Cattell (1971), has at its highest level a distinction between fluid-analytic intelligence (G_f) and crystallized-verbal

Table 1

Correlations among Aptitudes Reported by Peterson (1976), Porteus (1976), and Rutkowski and Domino (1975).

	Peterson Experimental Students (N=94)				Porteus Experimental Students (N=56)			
Aptitudes	G_c	A_i	A_c	A_x	G_c	A_i	A_c	A_x
G_c	.94[a]	.47	.34	-.10	.63	.33	.11	-.10
A_i		.48	.29	-.16		.62	.07	-.04
A_c			.65	-.45			.70	-.23
A_x				.84				.81
Rutkowski-Domino College Frosh (N=201)								
Delay Avoidance - SSHA[b]					-.02	.07	.48	.01[c]
Work Methods - SSHA					.32	.30	.41	-.31
Teacher Approval - SSHA					.17	.42	.41	-.12
Education Acceptance - SSHA					.10	.26	.49	-.12

Notes:

a Coefficients in the main diagonals are reliability estimates given by the respective authors. Porteus's coefficient for G_c is the communality from her factor analysis, and is thus an underestimate.

b The four scales of the Survey of Study Habits and Attitudes.

c Estimated as the reflected average correlations between SSHA scales and the five CPI scales with highest loading on Factor II, "Poise and Self-assurance."

intelligence (G_c). After decades of factor analytic debate, this or related hierarchical views have finally become popular because they fit existing data rather well while offering some hope of parsimony. More specialized distinctions among abilities, such as Guilford's (1967), can be fit into the more specific levels of the hierarchy. Horn (1976) summarized much recent correlational research under these rubrics, adding spatial visualization ability (G_v) to the general level. The provisional use here of this three-sided concept of G seems justified.

Cronbach and Snow (in press) adopted a hierarchical view in attempting to make sense of the large number of ATI studies that rest on measures of one or another of these constructs. Unfortunately no large-scale ATI studies attack this complex squarely. In most past research, ATI must be attributed to an undifferentiated G. And studies based on more specialized ability constructs cannot be understood as such, in the hierarchical view, unless they rule out interpretation in terms of G by including general measures. Few do. Thus, the present section must rely mainly on a scattering of small-scale instructional experiments that are impossible to integrate in a way comparable to the studies on the $A_iA_cA_x$ complex.

G has been the most widely studied aptitude construct in instructional research and thus has produced the most ATI. Table 2 attempts a summary of the kinds of treatments that seemed to alter the relation of G to outcome most substantially, based on the Cronbach-Snow review. The general hypothesis derived from this review is stated by Items 1) and 11) of the table; instructional treatments differ in the information processing burdens they place on, or remove from, the responsibility of the learner, and the regression slopes of cognitive outcomes on G become steeper or shallower accordingly. It seems that the more learners are required to puzzle things out for themselves, to organize their own study and build their own comprehension, the more able learners can capitalize on their strengths profitably. As instructional treatments are arranged to relieve learners from difficult reading, analyzing complex concepts, and building their own cognitive structures, the more such treatments seem to compensate for or circumvent, less able learners' weaknesses. These latter treatments help Low G students; the High G students may or may not do well in them. Often, the impression from such studies is that High G students can do well enough no matter what treatment is applied.

Table 2 here

But this is not always the case. Item 16) in Table 2 suggests that some kinds of treatment-supplied cognitive models can actually be harmful to High G students. Whether such an effect is widespread, and whether it holds for G_c, G_f, and G_v, or for only one or two of these constructs, are important research questions. G_c and G_v are easily separated correlationally. (This is essentially the distinction between v:ed and k:m consistently reported in British factor analytic studies.) But G_f is not readily separable from the other two in this way. Experimental manipulations that could make such distinctions would sharpen construct validity, especially of G_f, as well as providing process hypotheses about ATI.

Two lines of research relating to this question, by Salomon and by Greeno and Mayer, were summarized by Cronbach and Snow. They are treated here briefly, followed by the most recent studies by Mayer and Greeno, along with scattered other work.

An early study by Salomon (1968) used a commercial film to give either "cue attendance" or "hypothesis generation" training to student teachers. The former asked the students to notice many stimulus details in the film; the latter asked them to produce possible themes for the story depicted. The film scenes were shown in scrambled order to some groups. A posttest was scored for information seeking (questions asked in a problem-solving transfer task about a curriculum organization problem). G_c (GRE-V) related to outcome positively after hypothesis generation training, but negatively after cue attendance training. Apparently, Low G_c students were helped by practice in attending to details, but High G_c students were hurt (or bored) by this. Hypothesis generation may have required more synthesizing, which perhaps overloaded Low G_c students while challenging High G_c students.

Salomon (1974) then investigated the effects of modeling a skill process by film. Two studies concerned the cue-attendance skill, defined in terms of both quantity of details noticed in visual displays and organization of spatial scanning (perhaps an aspect of G_f and G_v). The visual displays were complex paintings. A cue attendance measure served as both pre and posttest, yielding quantity and organization scores. Another posttest measured information search in problem-solving. In one treatment condition, a film demonstrated the process of zooming in on closeups of specific details one after another in sequence.

Table 2

A Summary of ATI Hypotheses Relating General Ability (G) to Various Instructional Treatments (T)

When T is described as...	... the expected results for students described on G as Low and High are...
1) placing burdens of information processing on learners,	-[a] and + respectively
2) using elaborate or unusual explanations,	- " + "
3) a "new" curriculum,	- " + "
4) including discovery or inquiry methods,	- " + "
5) encouraging learner self-direction,	- " + "
6) relatively unstructured or permissive,	- " + "
7) relying heavily on verbiage,	- " + "
8) rapidly paced,	- " + "
9) giving minimal essentials by PI[b], for learners to elaborate	- " + "
10) giving advance organizers on difficult material,	- " + "
11) relieving learners of information processing burdens,	+ " 0 "
12) giving all essentials by PI,	+ " 0 "
13) simplifying or breaking down the task,	+ " 0 "
14) providing redundant text,	+ " 0 "
15) substituting other media for verbiage,	+ " 0 "
16) using simplified demonstrations, models, or simulations,	+ " - "
17) varying the format of PI,	inconsistent
18) including inserted questions,	inconsistent
19) using diagrammatic or pictorial presentation,	inconsistent
20) based on specialized film or TV.	inconsistent

[a]For "-" read "poor"; for "+" read "good"; for "0" read "uncertain".

[b]PI = Programmed Instruction.

Another treatment merely gave practice (called "activation") in noticing and listing details. ATI were marked; the film model helped students initially low in cue attendance but disrupted already skilled students, who did better with simple activation. The same result occurred for the problem-solving outcome in the first study but not in the second (although it was measured differently here). Also, in the second study, measures of G_f (Embedded Figures) and G_c were included, and these gave the same ATI results as did the cue attendance pretest. It was hypothesized that giving High G students a processing model which they do not need causes interference with their own idiosyncratic processing. Also noted was the finding that scanning by able students appeared more organized after model training; they seemed to organize their search according to the model at the expense of listing details, while less able students showed the opposite pattern.

A final experiment used a film model to demonstrate the sort of folding and unfolding of three dimensional objects often found in measures of G_v. Pre and posttests were two kinds of Thurstone G_v tests. Two other pre measures represented G_c (grade point average in language and mathematics courses). Modeling seemed to improve G_v for all students, with no ATI. But language grade interacted; it was positively related to the G_v outcome measure in a control group and negatively related in the model group. Salomon suggested that the normally logical, sequential processing of high verbal students might be disrupted by attempts to use the spatial operations given by the model. It is not clear why mathematics grade would not have given the same result, except that higher-level mathematical ability (and presumably also mathematics achievement) has been found related to both G_c and G_v (i.e., it may be close to G_f). For a review of G_c-mathematics relations, see Aiken (1971).

The work of Greeno and Mayer serves several task analytic purposes in instructional research in addition to the interest at hand. In one study, Mayer and Greeno (1972) compared a rational-conceptual method with an algorithmic formula method for teaching binomial probability by CAI to college students. They obtained ATI using a pretest on probability and permutations, but not with a mathematical ability test or with self-reported SAT-Q score as aptitudes. Students with high pretest scores did better with concepts, and worse with the algorithms than did students with low pretest scores. Apparently, the more rational treatment required knowledge of prerequisite probability concepts and permutations, hence was best for Highs. The algorithmic approach did not require

these prior structures and so was better for Lows. It can be hypothesized further that Highs did poorly with algorithms because the latter did not fit the modes of operation preferred by these students, or because they were not rationalized as meaningful in relation to students' prior concepts. As in some other studies, a kind of learning hierarchy may be implied by this ATI finding, where alternative treatments are best at different stages of knowledge acquisition.

The authors went on to suggest that the two treatments produce different structural outcomes. The conceptual condition evidently produced a cognitive structure with good external connectedness (with components easily separable for use in problem solving). The formula condition produced good internal connectedness (easier application of the concept as a whole).

Egan and Greeno (1973) then compared a rule method with a discovery method in a similar situation. Again, specific pretests and SAT-Q were the aptitudes. The rule group was taught the formula for combining binomial probabilities while the discovery group had to infer it from examples. A similar, though ordinal, ATI occurred. The posttest had strong relation to aptitude (both specific and general) in the discovery group and almost none in the rule group, so discovery was particularly disadvantageous for Lows. Again, structural differences in outcome were noted. Discovery learning was characterized as producing external connectedness similar to the outcome for conceptual learning in the earlier study. This was interpreted as supporting Gagné's view that the learning of meaningful concepts and the discovery of principles result in similar outcomes. Rule learning gave better internal connectedness, as did the earlier formula treatment.

A further distinctive feature of this study was the use of a multifaceted transfer measure, with problems stated in both words and symbols, and of both near and far transfer (Luchins) types. There were interactions between aptitude and type of problem. The combination of aptitudes related to performance on all kinds of problems, but related most strongly to the far transfer-Luchins problems. As many other studies have suggested, regression slopes were steeper in discovery than in rule conditions.

A smaller study was also reported here using programmed booklets instead of CAI. Specific pretest gave roughly similar ATI trends as in the main study, but the more general aptitude tests did not interact.

Mayer, Stiehl, and Greeno (1975) have produced extensive further work along these lines. There were four experiments, all with college students.

In the first study (N=44), students received CAI instruction in binomial probability using a formula or a general concepts program. A crossed treatment variable concerned whether or not questions were interspersed in the text. Again, a faceted posttest allowed qualitative distinctions in learning. Aptitudes were self-reported SAT-Q, two specific pretests on probability and arithmetic concepts, and a permutations test. There were no main effects for treatment. Dividing students into Low, Medium, and High on each aptitude in turn, ATI results quite like those of Egan and Greeno were obtained. On the probability and permutations pretests, the formula condition was best for Lows while the general concepts program was best for Highs. The arithmetic pretest gave the same pattern for Highs but treatments did not differ for Lows. Self-reported SAT-Q did not interact. Beyond its value as replication, this study is important in another respect; it probed further into the possibility that aptitude-treatment-posttest interactions (ATPI) might occur. If AT combinations resulted in qualitatively different learning outcomes, then aptitude would be relevant not only to amount learned in a treatment but also to the kind of learning produced by a treatment. However, ATPI were for the most part non-significant. Some trends suggested, for example, that among medium aptitude students (on the probability pretest), the formula treatment produced higher performance on questions stated in formula form while the general concepts treatment was better for questions stated as stories, with no such differences among low and high aptitude students. Had such patterns occurred at other levels, or more importantly, had different patterns occurred at different aptitude levels, a whole series of important further questions would arise in ATI work. Faceted posttests of the sort used by Greeno, Mayer, and their coworkers should become routine in instructional research, including ATI studies. The ATPI possibility remains an important question to ask, especially in studies of longer duration.

The other three studies of Mayer, Stiehl, and Greeno dealt with manipulated prior experience rather than its measurement via aptitude tests. In one (N = 90), students received no introductory problems, some introductory problems but no feedback, or problems with feedback and correction. The same instructional treatments as above were then applied. Again, main effects were not significant, but the interaction of prior experience and treatment was significant and similar in some respects to the earlier ATI findings.

Students without prior experience did better with the formula treatment. The treatments did not appear to differ for experienced students. There were also some experience-treatment-posttest interactions, indicating that the nature of instruction made some qualitative difference in outcome for students with prior experience, but not for those without it. The other studies (N = 50; 40) gave preinstruction in formula writing or in the meaning of concepts, or both or neither, and then placed all students in a common instructional sequence. This was designed to show the difficulty of different steps in learning. This kind of pretraining manipulation seemed to produce a set to learn particular aspects of the material. Formula experience, for example, produced strong performance on formula questions but poor performance on story problems, while no experience or experience in both treatments gave better results than did formula experience, particularly on story problems. General concept experience alone was best for story problems.

From all this, the authors draw the following three-part conclusion:

1) ATI results suggest that formula instruction should be given to low aptitude students, to increase total amount learned.

2) If, however, ATI arises from fairly specific kinds of prior knowledge, a better policy would be to provide low students with training to remedy this.

3) This is so because, while the formula treatment produces skillful performance, it does not produce more general understanding, and so does not prepare the low aptitude student for later learning.

This reasoning is sensible enough as far as it goes. It emphasizes the important role that prior achievement may play as aptitude, a point also made by Tobias (1976) in a review of other such studies. And, it harks back to the Brownell and Moser (1949) results (discussed by Cronbach and Snow). In a multi-school comparison of rote vs. meaningful arithmetic instruction, they found that some third-graders seemed unable to profit from explanations. Their prior rote training had apparently developed an inaptitude for later meaningful instruction. Thus, short-run, expedient decisions about alternative treatments may be detrimental; one needs to take a longer view.

But a longer view involves at least three other issues, and the present results do not address them. First, learning time in the Mayer and Greeno experiments is measured in minutes while school learning is measured in months. In school learning, it is important to ask: How much prior training

will be required for students low initially on a specific achievement pretest? How far will this training carry them into the next block of instruction? How far behind other students will they be as a result? When specific inaptitudes can be made up easily, without other costs, that is clearly the best line for the educator to take in applying ATI data. But if remedial instruction becomes the predominant mode for some students, other kinds of alternate treatments need to be considered.

Second, and perhaps more importantly, one should ask: Do not both high and low students initially need ultimately to have both kinds of instruction? An old study by Edgerton (1956; see Cronbach and Snow) suggested just this. A 14-week course for weather observers was taught in two alternative sequences; theoretical explanations first, then practical techniques, or the opposite. On a faceted posttest much like those used by Greeno and Mayer, items could be divided into "How" questions and "Why" questions. On "How" questions, the theory-first treatment was especially beneficial to students low on numerical ability (G_c?). This treatment had apparently given these students concepts that would help them when they reached computation. On "Why" questions, the techniques-first treatment was better for students low on reasoning ability (G_f?). They were allowed to master computation to help when conceptual reasoning was required. The implication is that complex instruction should start with the treatment that avoids each student's weakness, even if both treatments need ultimately to be given. Another Edgerton (1958) study compared rote vs. meaningful explanation in an aviation mechanics course. As it turned out, G_c related more strongly to performance in the rote treatment than in the meaningful treatment. Apparently, explanation helped Lows while Highs generated their own explanations.

Third, tests of specific prior knowledge must be connected to the more general fabric of individual differences. One cannot study mathematical ability and achievement in isolation just because it appears to be more directly relevant. Ordinarily, all measures of prior knowledge including specific pretests can be thought of as representing G_c. It would be convenient to do so here, but for the apparent fact that the ATI findings came mostly from the specific pretests rather than the general tests. However, the four aptitude tests used were undoubtedly intercorrelated and one can speculate that ATI was mainly associated with the principal component of these. Hence the simplest interpretation of the results should rest on

$G_c G_f$, until future research shows conclusively that general constructs cannot account for specific ATI. As noted earlier, mathematical abilities often seem to combine G_c, G_f, and G_v.

The most recent work in this series comes from Mayer (1975). He taught computer programming to undergraduates using a diagrammatic model of a computer expressed in familiar terms as one treatment, and a condition using a rule text with examples but no overarcing framework or model as the other. Three studies were reported, each using extremely brief instructional booklets.

The first study (N=86) compared two model and two rule treatments, one of each including an introduction to flowchart symbols. Aptitudes were SAT-Q, and tests on algebra computation, algebra word problems, card trick problems, and permutation ordering. The posttest required generation and interpretation of programs, with and without loops. A main effect favored the model. It was not clear how ATI was tested; presumably separate analyses treated aptitude as a two-level blocked factor, since later studies did this. Mayer reported there were no ATI. SAT-Q and the algebra tests did, however, give substantial correlations with various facets of the outcome measure.

In the second study (N=40), aptitude was represented in the design by SAT-Q, split into high and low ability blocks at score 560. Again, two rule and two model groups were compared, one of each including an introductory program. There were two program learning exercises. An additional outcome measure was a transfer set of programming exercises. There was no main effect due to the model, but Lows made fewer errors after instruction with the model while Highs were better off without the model. This pattern occurred in both learning exercises as well as the transfer exercises. (The trend was not statistically significant on transfer, but a small N design with aptitude blocked does not provide a powerful test.)

The third study (N = 56) used essentially the same design and measures but crossed model vs. rule text with two types of practice (interpretation vs. generation of programs). The same ATI was obtained. Also, the three-way interaction was noteworthy; while the rule treatment was best for Highs and the model treatment was best for Lows, in the generation practice condition, the rule treatment was best for both ability levels in interpretation practice.

Mayer concluded that high ability students may be hurt by the model because they already have their own idiosyncratic "models", in the form of rich prior knowledge. Low ability students seem to experience no such interference.

All these studies hint that treatments based on simplified models, algorithms, direct rules to follow, etc., may be detrimental to High G students, even while they help Low G students. This is a special case of the general hypothesis stated earlier about the locus of the information-processing burden, student or treatment. It is an especially interesting case because it implies negative effects on High G students. Some kind of interference phenomenon, operating between what the model tries to help students do and what students would ordinarily do on their own, seems to be responsible. But the sources of this interference cannot be pinned down more exactly at present. Nor is it clear that the phenomenon is more associated with G_c, as might have been hypothesized following one Salomon study. All of the studies show the effect with G_c. One of Salomon's studies failed to show it with G_v. But none of these studies effectively distinguishes G_c and G_f, or any of these general constructs from the special pretests preferred in some studies. Such distinctions will be important in future research; sharpening the aptitude constructs at work in ATI where Highs are hindered by a model or algorithm should help define the interference mechanisms that produce such results.

Two other studies bear more directly on the G_c G_f distinction. But they fail to make it clear.

Skanes et al (1974) used a brief treatment to explain and give practice in solving problems. Over 2000 students in Grades 5 to 9 participated. Some students received practice on letter series, some on number series. These groups were further subdivided, to form pretest and no-pretest groups. A fifth group served as control. Pretest and posttest were forms on Thurstone's Letter Series Test. Aptitudes were the Raven and Otis mental tests, the first of which is often considered as representing G_f, the second as G or G_c. Multiple regression analysis showed two significant ATI. It appeared that students high on Otis were helped by having the pretest while students low on Otis were better off without it. Also, students low on Raven were helped by direct training on letters while those high on Raven derived more benefit from the indirect training using numbers. The analysis did not distinguish G_f and G_c satisfactorily; Otis and Raven were undoubtedly correlated. Thus the results must be attributed for the time being to an undifferentiated G. If the direct vs. indirect variation could be shown to interact specifically with G_f, it would imply that the information-processing burden hypothesis discussed previously applies to G_f as well as to G_c.

Christ-Whitzel and Hawley-Winne (1976) conducted a year-long evaluation of individually prescribed instruction (IPI) in mathematics, using 124 sixth graders. Treatments were IPI, a mathematics management and support system which related different objectives to available curriculum materials, and provided individual and small-group as well as whole-class teaching, and a traditional basal text with whole class teaching. Each treatment was applied in three classes of a different school. Two forms of the Comprehensive Test of Basic Skills served as pre- and posttest, administered in Spring of the prior year and in Spring of the treatment year. This gave a total score and subscores for computation, concepts, and applications. While these were used as separate outcome measures, the total score alone was used as aptitude. Also, an IQ score available at pretest correlated highly with the pretest scores, so it was not included. The pretest total score can probably be thought of as G_c. Other aptitudes were Embedded Figures (G_f) and questionnaires giving several self-concept and locus-of-control scores.

Multiple regression analyses showed G_c accounting for most of the variance in each outcome measure, with G_f adding significantly each time. Treatment main effects showed traditional instruction highest and IPI lowest, but these differences also existed at pretest so no average conclusion is justified. ATI appeared for both G_c and G_f, and they seemed to run in opposite directions. For total outcome and computation, Low G_c students were better served by IPI while High G_c students did better in traditional classrooms. After extracting this interaction, however, High G_f students seemed to do better in IPI! For the other two outcomes these aptitudes continued to show main effects. The ATI trends were as before but very faint. Self concept showed one other ATI; on the concepts subtest, the curriculum management treatment was best for Highs, while the other treatments were best for Lows.

The analysis is an interesting demonstration of ATI used in evaluation, but it does not go far enough for the present purposes. Without knowing correlations among all the pretest measures, or sharply distinguishing G_c and G_f as predictors, the implication is not fully tested. An important hypothesis, however, for re-analysis and for further research, is that individualization (via IPI) helps those who have not learned well in prior instruction but who have high fluid ability ("underachievers"?); it is less effective for those high in prior

achievement and/or low in fluid ability. This result would replicate the old Anderson (1941) study, which gave a similar implication in a comparison of drill vs. meaningful instruction. As will be clear in the methodology section to come, however, the Anderson finding does not withstand reanalysis.

This section can be closed with a series of studies dealing with the deceptively simple distinction between G_c and G_v. The most popular special ability hypothesis in ATI research has been that a verbal vs. visual distinction among abilities should interact with similar distinctions in instructional materials. The simplest hypothesis was that High G_v-Low G_c students would excel in visual simplest hypothesis was that High G_v-Low G_c students would excel in visual treatments while Low G_v-High G_c students needed verbal treatments. The Cronbach-Snow review showed much inconsistency in attempts to verify this (see Table 2) and suggested that the hypothesis might often be stated in the opposite way. Visual treatments might at times compensate for Low G_v; verbal treatments might compensate for Low G_c. Difficulty of the material would be one important moderating factor in such hypotheses. Also, general vs. special ability interpretations needed to be clarified.

Gustafsson (1974; 1976) has pursued this problem in three studies done in Sweden. The work includes a thorough rethinking of ATI in this domain, and a methodological comparison of various regression and blocked anova approaches. The first study (N=316 seventh graders) used a 9-page text on the polar lights phenomenon as a verbal treatment, with a slightly reduced text plus 17 illustrations (16 pages) as a pictorial treatment. Aptitudes were Opposites (G_c), Number Series (G_f?), and Metal Folding (G_v), plus some interest scales. Outcomes were short-answer and essay tests administered directly after instruction. G_v showed stronger relation to achievement in the pictorial treatment, while G_c was more highly related to performance in the verbal treatment, particularly on the essay outcome. Number Series gave results rather like those of G_c. This finding, then, demonstrates the traditional expectation. But some special concerns about the study led to a reanalysis (summarized in Gustafsson, 1976). This allowed quadratic terms and also checked some effects by blocking on aptitude. The results for the essay test remained essentially unchanged; verbal instruction was best for students low on G_v and high on G_c while pictorial instruction was best for students high on G_v and low on G_c. But this effect had to be attributed in part to interference from some association tests which had

been administered during instruction in some groups; they apparently had had a negative effect on the reading strategies of High G_c-Low G_v students and could have accentuated the ATI. Results for other criteria were modified in a complex way. Apparently, verbal instruction served High G_v-Low G_f students best on the short answer test. A complex three-way interaction among the aptitudes was regarded by the author as due to measurement problems. Finally, for those short-answer items that dealt directly with the illustrated material, pictorial instruction was best for Low G_c students and for Medium G_c students who were high on G_v. This was apparently regarded as the most trustworthy of the results.

Gustaffson's (1976) second study (N=201 seventh graders) sought to replicate these findings using the same instructional material and the same three aptitudes. The short-answer test was again divided into items bearing on illustrated vs. nonillustrated material. This plus the essay test were administered as immediate criteria and again eight days later as retention measures. The pictorial treatment produced somewhat higher achievement on all criteria. Overall, ATI were not significant. However, the score for illustrated short-answer items gave steeper regression slopes on G_c in the verbal treatment than in the pictorial treatment, so students low on G_c did best with pictures. This is consistent with the reanalyzed results of the first study. The effect reached borderline significance at retention. As in the first study, G_v showed a tendency to augment this effect, so that the best performances in the pictorial treatment among students low on G_c occurred among those also high on G_v. The results for essay were mixed and failed to conform to earlier findings.

A third study (N=229) used new instructional materials on the heart and blood circulation system. Again, brief text materials were presented in illustrated and unillustrated versions. In contrast to previous studies, the subject-matter was especially designed to be "spatially demanding", on the hypothesis that Low G_v students would thus profit most from pictures. Immediate outcome measures yielded three scores: verbal content achievement, verbal achievment on spatially demanding content (called "spatial"), and pictorial content achievement. Aptitude tests used and factor analytic results for them are shown in Table 3. Gustafsson chose to use simple aptitude scores in some ATI analyses, factor scores in others, and some combinations of simple scores in still others.

- - - - - - - - - -

Table 3 about here

- - - - - - - - - -

Table 3
Varimax Rotation Results from Factor Analyses Reported by Gustaffson (1976), Satterly (1976), and Das and Molloy (1975)

	Tests[a]	G_c	G_f	G_v	PA	PS	M
Gustaffson	Opposites	.83*		.16	.07		
Seventh Graders	Reading Speed	.80*		-.04	.15		
(N=229)	Number Series	.66*		.37*	.09		
	Metal Folding	.42*		.53*	.25		
	Figures	.15		.82*	.01		
	Hands	.03		.77*	.01		
	Word Pair Associates	.21		-.01	.88*		
	Picture Pair Associates	.05		.12	90*		
Satterly	Verbal Intelligence	.76*	.26	-.08		.05	
Fifth-Sixth Graders	English Comprehension	.83*	-.07	.27		.03	
(N=201)	Picture Vocabulary	.79*	-.01	.38*		.16	
	Mathematics Attainment	.67*	.30*	.09		.04	
	Embedded Figures	.31*	.77*	-.04		.13	
	Haptic Perception	.32*	.48*	.23		.18	
	Analytic Preference	-.04	.64*	.10		-.18	
	Gottschaldt Figures	-.02	.48*	52*		.25	
	Spatial Judgment	.34*	.13	.65*		.25	
	Spatial Ability	.24	.03	.87*		-.08	
	Perceptual Speed	.11	.01	.10		.93*	
Das and Molloy[b]	Verbal IQ	.59*	.08			-.16	-.12
Fourth Graders	Backward Digit Span	.76*	.28			-.03	.14
(N=60)	Bridge Task	-.54*	.34*			.04	-.18
	Performance IQ	.22	.57*			.08	-.23
	Raven Matrices	-.01	.86*			-.02	-.02
	Figure Copying	.07	.75*			.18	.00
	Memory for Designs (Errors)	-.03	-.70*			.03	.07
	Cross-Modal Coding	.01	.64*			-.11	.14
	Word Reading	.03	.04			.84*	-.14
	Color Naming	-.16	.06			.78*	.01
	Forward Digit Span	-.14	.04			.00	.80*
	Visual Short Term Memory	.14	-.08			-.22	.67*
	Serial Recall	.05	-.07			.01	.93*
	Free Recall	.08	.00			-.01	.90*

Notes:

- * Factor loading exceeds .30.
- a Published Test Lists have been reordered here to place similar tests together.
- b In the Das-Molloy List, two variables defining an SES factor have been omitted.
- PA = Paired associates learning factor.
- PS = Perceptual speed factor.
- M = Short term memory or memory span factor.

Various analyses failed to show substantial ATI. There was a tendency toward curvilinear regression, with the pictorial treatment giving best results on the spatial criterion for students scoring either low or high on the Figures Test (G_v). But this trend also seemed present for the verbal criterion! In general, Gustafsson concluded that for verbal and spatial criteria there were no ATI. For the pictorial criterion, there was ordinal interaction; the pictorial treatment was best for everyone but least so for students high on G_c and low on G_v. This is contrary to the hypothesis, and it does not seem to match up with the results of the earlier studies. The conflicting character of these studies is typical of research on the verbal vs. visual contrast, as noted earlier.

Table 3 has been constructed to illustrate a point also alluded to by Gustafsson. Number Series and Metal Folding had correlations with members of both sets of tests, and the factor analysis split them between G_c and G_v. This seems to be a typical result, not specific to these data. One might conclude that performance on these tests requires both G_c and G_v, or that their items can be solved using either ability, or that they represent abilities involved in the other two abilities, namely G or G_f. This latter possibility is often lost in such analyses; it might have been made to appear in the Gustafsson data by the addition of measures such as the Raven or Wechsler Performance IQ, or by a factor solution closer to the unrotated axes than the varimax procedure allows.

Data from two other studies are shown in Table 3. Satterly (1976 sought to distinguish the cognitive style called "field independence" from G and then relate it to mathematics achievement. This style construct has process implications, so understanding might be advanced thereby. But his choice of reference measures produced only G_c and G_v, so his style factor cannot be distinguished from G_f and is so labeled here. As in the Gustafsson data, several measures including the mathematics score may represent either or both of two abilities, or a more basic one; measures such as Raven or Wechsler Performance IQ are again needed. The Das-Molloy research (1975; see also Das, Kirby, & Jarman, 1975) represents a unique attempt to reinterpret ability factors in process terms, using Luria's distinction between simultaneous and successive modes of information integration in cognition. They interpret what is here labeled G_f as simultaneous because it includes several tasks that involve spatial-constructive or multisensory performances, in addition to the traditional G_f markers. This is a provocative suggestion for research, but not an argument for creating new factor labels. The Bridge Task in particular may signal an important direction

for such research, with its negative loading on G_c. But what is here called G_c was interpreted by Das and Molloy as proficiency in spatial-imagery. There is no justification for this in the data, and no G_v measures were included in the battery. The factor called "successive" must be listed here as short term memory; its tasks are common memory measures, which are admittedly serial in character.

Thus, Table 3 illustrates the problems endemic to much of differential psychology. Aptitude factors can be extremely useful as reference points for sorting individual differences. But alone they can never be interpreted as unities, or yield process theories, or sustain new interpretations. Correlational analysis alone will never solve these problems. What is required are process descriptions of how students solve the problems represented in different tests, and this will require introspective interviews as well as much experimental analysis. It may be found, for example, that Metal Folding problems are solved by some students using visualized rotations, and by others using analytic reasoning. This is the implication of early research by French (1965). Until aptitudes as well as treatments can be described in such terms, inconsistent ATI results, and consistent results as well, will not be satisfactorily understood.

In sum, work on the $G_cG_fG_v$ constructs is insufficient to justify their distinction in ATI research. The recommendation is not to dispense with the distinction, but rather to sharpen the analytic attacks in both instructional and laboratory research. No ATI study has yet measured the three adequately. At the same time, the correlates of each should be sought in the broader domain of differential psychology. No research that includes a mental measure studies that variable in isolation. The personal, motivational, and social correlates of these constructs may help to distinguish them from one another and to connect them to other instructionally relevant complexes, such as $A_iA_cA_x$. A long line of research has sought to accomplish this for the verbal vs. quantitative ability distinction (see McCarthy, 1975). The same is required for $G_cG_fG_v$.

In the meantime, the interpretation of most cognitive tests in ATI rests on G. And, for the sake of parsimony, any study investigating a more specialized aptitude must show its separation from G empirically. The only other cognitive ability constructs thus distinguishable, that also have had some place in ATI research, are the other factors listed in Table 3; memory measures, perceptual speed, and some specific learning performance measures. Memory tests, in particular, have occasionally shown distinct interactions. These will be touched on in the final section of this review.

Methodological Developments

Cronbach and Snow argued that tests of ATI hypotheses would have to be designed as large-scale, long-duration, real school studies. Only in this way could stability of aptitude and treatment effects be achieved, with sample sizes sufficiently large for powerful statistical tests, in settings likely to lead to practically important conclusions. Studies on this scale are expensive and time-consuming, so they occur relatively infrequently. But they are increasingly seen as worth the effort; the Peterson and Porteus studies step in this direction. Also, the incorporation of an ATI viewpoint into large-scale evaluation studies has increased, and secondary analyses of data from such studies has helped to test some ATI hypotheses at minimal extra expense.

In all ATI work, multiple regression methods have been advocated as the most general and appropriate approach to data analysis. With new texts now available (see Kerlinger and Pedhazur, 1973; Amick and Walberg, 1975; Borich, Godbout, and Wunderlich, 1976), multiple regression methodology is coming into routine use.

A new look at regression methodology in large scale studies, however, has introduced some new problems with potentially profound implications for future research. These problems arise from methodological concerns, but they have central substantive importance. They are introduced in this section, to spur interest and further development, but they cannot be resolved here.

Between-class and within-class regression components. The large, real school studies advocated for ATI research must inevitably include multiple classrooms, with classes assigned to alternative instructional treatments. There are then individuals nested within classes nested within treatments. Virtually all research of this character has ignored class groupings, producing pooled regressions of outcome on aptitude for individuals under separate treatments. But this assumes that outcome is a function of individual aptitude only, while class composition on aptitude, class average aptitude, and individual teachers have no effect. Theoretically, one could posit two kinds of aptitude effects in addition to individual effects: comparative effects, wherein a student's standing on aptitude in a particular class is important, if e.g., the teacher pays primary attention to those students above average on some aptitude in that class; and class effects, wherein class average aptitude is important,

if, e.g., the teacher adopts a fast pace in a class with a high average on some aptitude. The pooled individual regressions for treatments combine, and mask, both comparative effects and class-level effects. Clearly, a methodology for multi-class studies is required that separates between- and within-class regressions, so that alternative causal explanations for ATI in any given instance can be examined. Furthermore, it becomes clear that single-class experiments are unable to distinguish among these alternative sources of ATI, because between-class effects cannot be detected here. In such studies, routine interpretation of ATI as bearing on individual phenomena is unjustified.

Cronbach and Webb (1975) reanalyzed data from an early study by Anderson (1941) that had seemingly shown a marked and important ATI. In the Anderson study, 18 fourth grade classes were assigned to drill vs. meaningful instruction in arithmetic. Using an achievement pretest and a general ability measure as aptitudes, and an achievement posttest as outcome, Anderson computed regression slopes at the individual level ignoring class membership, and obtained ATI. Students high on prior achievement and low on ability ("overachievers") did best in drill, while students low on prior achievement and high on ability ("underachievers") did best with meaningful instruction. This result would be important still today, implying as it does that students *high on* G_c *and low* on G_f should be given drill instruction, a method in which they had already displayed progress (drill being prevalent in arithmetic instruction in the late 1930's), while students high on G_f and low on G_c should be switched to the more innovative, meaningful instruction. The Anderson study is one of the few large, real school studies bearing on the hypotheses discussed earlier in the section on the $G_cG_fG_v$ complex.

The Cronbach-Webb reanalysis (N=434) estimated between-class regressions to test class-level effects, and within-class regressions to examine comparative effects. This analysis does not separate these and individual effects completely, but does give us hints as to which effects are likely to be worth further attention. In brief, the individual ATI implied by the Anderson analysis vanished. The class-by-class regressions varied greatly, so that the pooled within-class within-treatment regression slopes gave only a fraction of the trend reported by Anderson's analysis. At the between-class level, the small number of class means and some statistical anomalies of the Anderson sample made a conclusion for or against ATI unjustified. The reanalysis thus wipes away the Anderson

study as support for an ATI hypothesis involving G_c and G_f. More importantly, it cues present researchers to attend to complexities in ATI research, and in all classroom research, not dreamed of only a short time ago.

A contemporary demonstration of this complexity comes from a study by Greene (1976). She investigated the effect of giving students their own choice about when, for how long, and in what sequence, they would work on booklet lessons in class. Nine fourth and fifth grade classes were divided randomly into choice and no-choice groups. The no-choice groups worked on booklets designed to teach the solving of Thurstone letter series problems under teacher direction and scheduling. The choice students could schedule themselves. Aptitudes included a pretest on letter series and Lorge-Thorndike Verbal IQ. Outcome included a posttest on the workbook and a measure of interest in doing further similar work. Various motivational measures served as both pre- and posttests. The study spanned four weeks of class time with N=165.

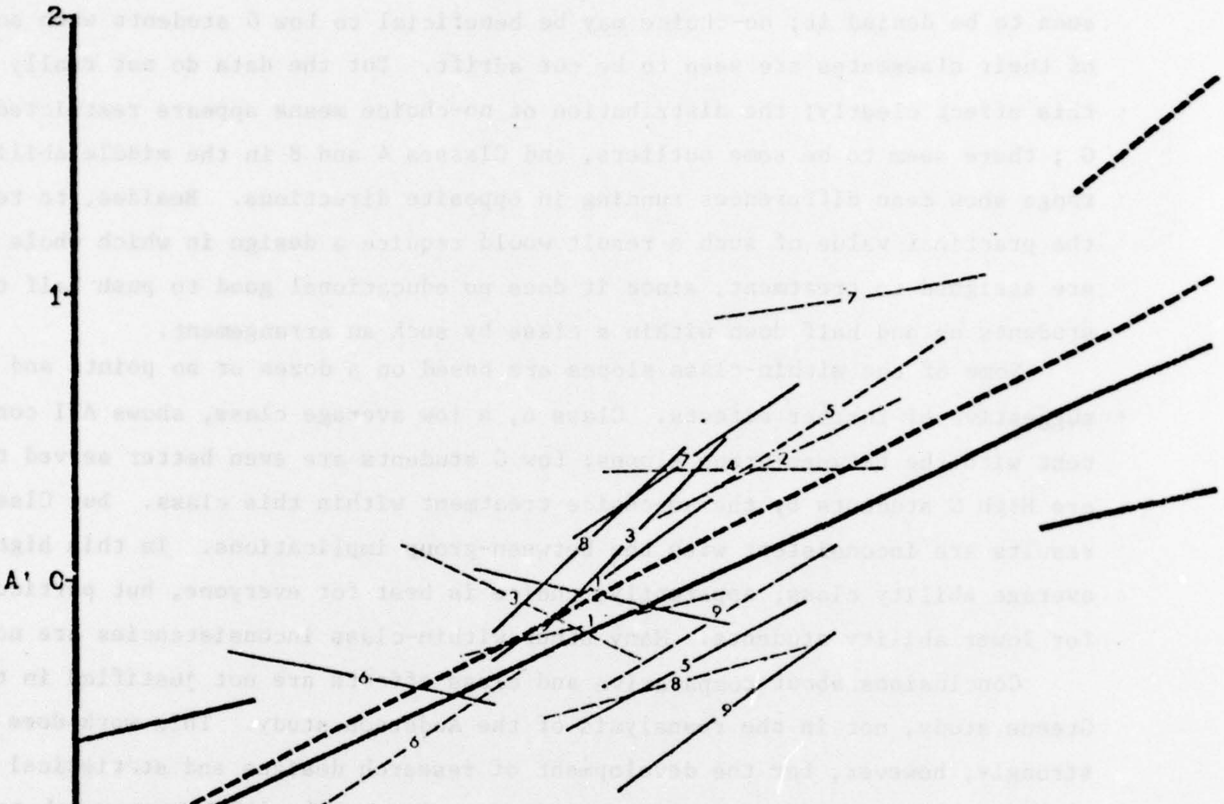
Figure 5 shows a representative result. Aptitude here is G combining Lorge-Thorndike and the letter series pretest. Outcome is a factor combining the letter series posttest and measures reflecting a student's confidence about school work (A'). The regression lines spanning the figure show the pooled individual slopes, for no-choice (solid) and choice (dashed) students. The regressions identified at the margins are between are between-class slopes fitted to the half-class means for each treatment. The thin lines show within half-class regressions, drawn to ± 1 S.D. around the numbered half-class means.

- - - - - - - - - -

Figure 5 about here

- - - - - - - - - - -

Only the pooled individual slopes are based on enough points to be trustworthy. Yet the impression of within-class heterogeneity, and of the divergence of between-class and pooled individual slopes is striking. Attending only to individual data, there is clearly no ATI. The between-class slopes, however, suggest a major effect. One might say that choice in organizing one's own work should be given to High G groups, while the teacher should do the organizing for Low G groups. But this may be so here only because both treatments exist within each class; the between-group slope for choice may appear in a steeper position, and that for no-choice in a shallower position, only because the social contrast is perceivable by the members of each half-class group in the classroom. It seems likely that such a social contrast would work in opposite directions in classes differing on average ability. That is, choice may



Pooled within no-choice treatment

Pooled within choice treatment

No-choice treatment within class i

Choice treatment within class i

Between class, no-choice

Between class, choice

Note: Sample range for G was -2.68 to 2.71
Sample range for A' was -2.47 to 1.82

Figure 5. Between-class, within-class, and pooled individual regression lines for two treatments using G as aptitude and A' as outcome (after Greene, 1976).

be relatively beneficial to High G students when some of their classmates are seen to be denied it; no-choice may be beneficial to Low G students when some of their classmates are seen to be cut adrift. But the data do not really show this effect clearly; the distribution of no-choice means appears restricted on G ; there seem to be some outliers, and Classes 4 and 8 in the middle ability range show mean differences running in opposite directions. Besides, to test the practical value of such a result would require a design in which whole classes are assigned to treatment, since it does no educational good to push half the students up and half down within a class by such an arrangement.

Some of the within-class slopes are based on a dozen or so points and are suggestive of further effects. Class 6, a low average class, shows ATI consistent with the between-group slopes; Low G students are even better served than are High G students by the no-choice treatment within this class. But Class 7 results are inconsistent with the between-group implications. In this high average ability class, apparently, choice is best for everyone, but particularly for lower ability students. Many other within-class inconsistencies are notable.

Conclusions about comparative and class effects are not justified in the Greene study, not in the reanalysis of the Anderson study. This work does argue strongly, however, for the development of research designs and statistical methodology capable of separating such effects. Cronbach's (1976) monograph takes a major step in this direction.

Outliers. Both the Cronbach-Webb treatment of the Anderson data, and the Greene data, show examples of how regression slopes can be profoundly affected by a few deviant data points. And the threat of outlier distortions increases as regression slopes are based on fewer cases. This is an inevitable problem; investigators will usually be limited to a dozen or so classes per treatment, with about 25 students per class. It is a problem as well in all ATI research, whether laboratory or school based. But what is an outlier? How can one decide that a particular data point is unnaturally deviant? Statistical criteria cannot serve alone. Psychological research requires substantive criteria. In any case, present methods are often specialized and incompletely developed. (See Klitgaard, 1976, for a useful summary.)

An example can be used to suggest some directions methodological development might take. It also serves to show why studies of long duration are required to understand the functioning of aptitude in school learning.

In an evaluation of a new curriculum for Dutch first-year (18 year old) medical students (Wijnen & Snow, 1975), aptitude-outcome relations were studied in each of ten four-week instructional blocks over the year. Instruction in each block proceeded by problem- or case-oriented, student-centered, small group discussion--a radical departure from the teacher-directed lecture and recitation to which Dutch, and most U.S. secondary students, have long been accustomed. While the obtained bivariate scatterplots would have been judged unusual by textbook standards, they were similar to some plots seen previously by the investigators, so deviant points could not be dismissed as chance outliers.

Figure 6a shows scores for the 49 students on a verbal reasoning pretest and an achievement test covering the first four-week block of instruction. The aptitude analysis began with the Verbal Analogies Test because it yielded the strongest simple relation to outcome among four ability tests and came closest of these conceptually to the definition of G. The regression line shown (corresponding to r =0.48) is fitted to all points taken together. The abnormality among high ability students is marked; it appears as if two bivariate distributions, running in contrary directions, have overlapped to reduce the overall regression.

- - - - - - - - -

Figure 6abc here

- - - - - - - - -

An informal analysis was constructed to identify underlying variables that might account for the results. The approach taken was related to one proposed by Marks (1964), and called by him "off-quadrant analysis". In effect, the criterion adopted to judge outliers was one of predictability; points whose deviance in an aptitude-outcome scatterplot can be predicted by psychological variables that make sense in context are <u>not</u> outliers. It was hoped that the method would be useful in identifying aptitude complexes worthy of more direct investigation.

The informal analysis worked with rules for fitting ellipses and partitioning the scatterplots into area groups, labelled by letter as shown in Figure 6a. Then, other aptitude information (in the form of personality scales administered at pretest) was used to identify variables associated with the B_1 to B_4 continuum, as well as with other area group contrasts.[2] Among other

[2]This procedure was developed by hand, using subjective judgment. A report on the method and analysis of the present data is in preparation. While formal statistical procedures are needed to replace this informal method, that step awaits the completion of later phases of this multi-year evaluation study, when replicates of the present scatterplot will be available. The purpose here is only to show the initial methodological point, not to draw conclusions.

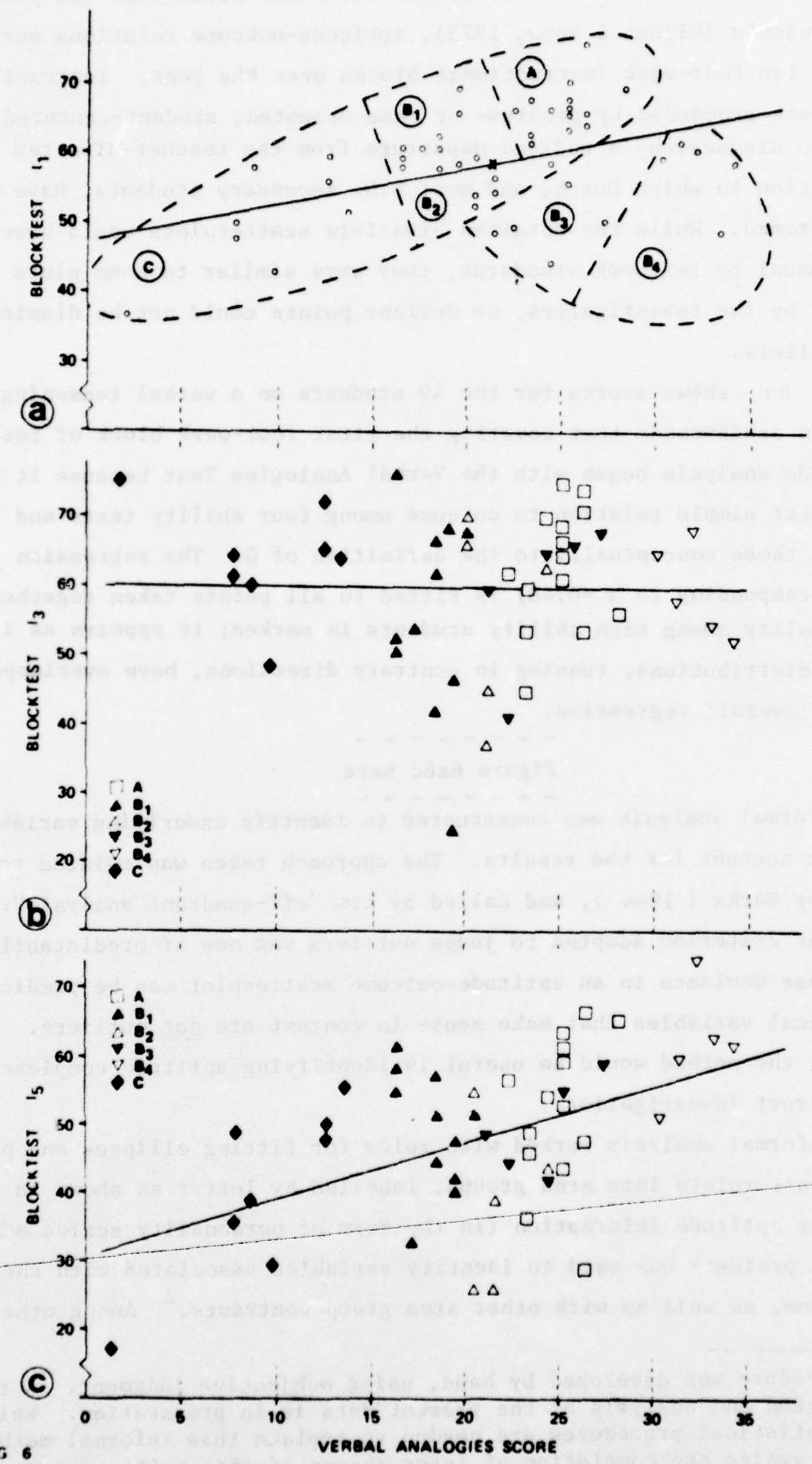


FIG 6

Figure 6. Scatterplots for three instructional blocks using analogies test scores as aptitude and block achievement tests I_1 (a), I_3 (b), and I_5 (c) as outcomes (after Wijnen & Snow, 1975).

findings, students in B_3 and B_4 described themselves are more independent and more motivated to achieve that did students of other groups. Together with A students, they also appeared less altruistic, less interpersonally oriented, and more task oriented than did students of other groups. B_4 students also showed the highest facilitating anxiety and the lowest debilitating anxiety scores. B_1 students were the most interpersonally oriented and the least task oriented, on average, of all the groups. The resulting hypothesis is that high ability students who are also highly motivated and oriented toward independent task activity do poorly in this instructional method, since it demands a large degree of cooperative interpersonal activity in learning. Middle ability students who value the interpersonal interaction perhaps more than the task activity do as well or better than many high ability students.

A further step of the analysis procedure was to follow these students through successive four-week blocks of instruction. Figure 6b and 6c show example scatterplots for two later blocks, in which students are identified by symbols corresponding to their area groups from the first block. The B_4 students appear gradually to adapt to the instructional situation, emerging at the top of the class by the 20th week (when content is also becoming more difficult and more biomedical in character). Some B_1, B_2, and A students seem to be falling out of the distribution across this same sequence; in effect, they are becoming outliers. Perhaps the novelty of the interpersonal instructional situation and their initial success in it is not sustaining. Thus, the aptitude complex that accounts for performance at an early stage shifts in later stages of instruction; what at first appeared to be an aptitude for this kind of instruction may appear later to be an inaptitude, and vice versa.

Although different measures were used in the Dutch situation than those of Peterson and Porteus, it is not hard to see the B_4 students as comparable to the High G, High A_i, High $A_c + A_i$, Low A_x students discussed earlier. Instruction was student-centered in the extreme, but it was socially-centered, which apparently conflicted at first with these student's independent, task-oriented styles of work, until they adapted to it or overcame it. This emphasizes again the importance of detailed analysis of aptitudes and treatments over time. It also provides an indirect argument in support of multivariate aptitude description even in educational experiments and evaluations not specifically concerned with aptitude.

Laboratory Studies of Aptitude Processes.

Before Binet's success in predicting school learning with a heterogeneous general test, Galton, McK. Cattell, and others had built measures of specific individual differences in sensory, perceptual, and memory skills. Their work was eclipsed in the ensuing development of mental testing, and the study of individual differences in learning-related processes has been a sporadic and isolated activity in psychology ever since. Cronbach (1957) and Glaser (1967) have traced much of this history. The Gagné (1967) symposium signaled the first significant return of laboratory research to this question. With the advent of information processing analyses of cognitive tasks at about the same time (Neisser, 1967), modern cognitive psychology has moved increasingly to the application of this view to the study of human intelligence (See, e.g., Resnick, 1975). This begs the more general issue of understanding individual differences in aptitude in terms of psychological processes. The need for this was noted earlier in connection with both Tables 1 and 3. The time finally seems ripe for the development of a coherent laboratory science of aptitude.

But this field is too new and far flung for systematic review at the present time. The best that can be done is to organize a list of findings and ideas that seem to point in useful directions. The list is little better than an annotated bibliography; it includes only studies appearing since the Gagné symposium that suggest relations between an aptitude construct relevant to this chapter and some kind of cognitive process measure. Primary emphasis is on analyses of $G_c G_f G_v$. Some ATI studies conducted in the laboratory spirit are also included. Left aside for the time being are experimental studies of A_x, many process analyses of laboratory tasks where aptitude constructs were not considered, and studies that attempt to train aptitude; these of course must ultimately be tied into the network.

An attempt has been made to order citations according to a "levels of processing" conception of cognition, from initial stimulus processing through reasoning and back to recall and response. It should be clear that this ordering is crude and at times arbitrary. Also to be noticed is a certain bidirectionality in the research; some investigators reason from aptitudes to processes, some from processes to aptitudes. This is as it should be; there is no a priori reason to consider one kind of measure more "basic" than another.

Initial stimulus processing. Trabasso and Bower (1968) proposed a model of attention in learning, including a parameter for stimulus sample size, i.e., the amount of stimulus information a subject takes in at a single glance. They regard this as a fairly consistent individual difference and suspect that it relates to such aptitudes as G, field independence (or G_f), and A_x. The Salomon work on cue attendance mentioned earlier may also fit here.

Boersma and his colleagues (Conklin, Muir, and Boersma, 1968; Boersma, et al, 1969) demonstrated that extreme groups on the Hidden Figures Test (G_f) differed in eye movement patterns during Hidden Figures item solution, and also on picture completion items. High G_f subjects attended to parts of pictures having higher information value than did Low G_f subjects. They also made more fixation shifts between simple and complex figures than Low G_f subjects. Males showed longer fixation times on the complex stimulus figure than did females, even though the latter took more total time per item.

A series of studies by Hettema (1967) suggested that individuals differing on three perceptual ability factors differed also in the character of their information reception processes. Dutch Army recruits were divided into profile groups using their factor scores on Flexibility of Closure, Speed of Closure, Optical Illusions, and Perceptual Speed. (The first two of these are Thurstone factors, but seemed to be subdivisions of G_f in this study). Then, in a Helson-style (1964) weight-judging experiment, it was shown that adaptation level formulae differed for subjects differing on perceptual factors; subjects high on Closure Flexibility were less affected by focal stimulation, those high on overcoming illusions were less affected by contextual stimulation, and those high on Closure Speed were less affected by residual stimulation. Other experiments gave the following results: subjects high on Closure Flexibility showed faster reaction times, especially under high focal uncertainty conditions; subjects high on Closure Speed were more apt to find short solutions to Luchins Jar problems, while subjects low on this factor more often opted for long solutions.

Day (1969; 1973a), using dichotic listening tests in which parts of words are fed to different ears, has found that subjects differ strikingly and stably in whether they are language-bound (fusing the two inputs to report a word) or stimulus-bound (reporting the stimuli without combining them.) The individual difference distribution was actually bimodal. The language-bound subjects appear to be poor judges of temporal order, while stimulus-bound subjects are

good at this. And, in memory span tests, language-bound subjects show marked serial position effect while stimulus-bound subjects show <u>no</u> serial position effect. In general, stimulus-bound subjects obtain higher memory span scores.

Another Day study (1973b) taught a secret language to subjects and then measured their ability to translate words into English. Secret languages use letter substitution or transformation rules of the kind found in pig latin. Language-bound subjects were shown to make far more errors in word translation than stimulus-bound subjects and to give more global, whole word translations as opposed to sequential, syllable-by-syllable, responses.

These and other data suggest qualitative differences in memory associated with this individual difference. Day believes that stimulus-bound subjects translate to visualizations (G_v?) while language-bound subjects do not, and that the former have a faster readout mechanism for transferring stimulus information from sensory register to short term storage.

This corresponds roughly to correlational studies by Seibert and Snow (1965; see also Snow & Seibert, 1966), where the erasure or backward-masking phenomenon in short term visual memory seemed to influence subjects differently depending on their perceptual and verbal ability profiles. Perceptual identification tests and G_c measures accounted for individual differences in erasure at different time delay intervals. For subjects high in both kinds of ability, the masking effect was less marked.

Rothkopf (1972) has studied individual differences in inspection rates during reading, changes in these rates as a function of inserted questions, and their relation to learning outcomes. These rates might also reflect stimulus sample size, speed of registration, or transfer to short term memory. In one study, he found that subjects with low performance on early segments of text <u>and</u> who accelerated their reading rate during this period, benefitted from questions inserted in later segments of text. Reading speed characteristics, in general, may provide a fundamental avenue for research relating aptitude and learning processes (see also Sticht, 1971). There is other evidence that inserted questions sometimes help low ability, particularly low memory-ability subjects, but can be detrimental to Highs. The evidence on inserted questions in relation to G is quite contradictory, however (see Table 2).

Sullivan, et al. (1971) studied phonics training for children who differed on a pretest of reading skills. He found the kind of ATI that suggests a hierarchical arrangement of discrimination skills. Training on single letters was best for low pretest students while training on letter combinations was best for students with more initial skill.

Cromer (1970) distinguished between "deficit" poor readers (those unable to identify words) and "difference" poor readers (those unable to integrate separate meanings in sentences), using vocabulary and reading score profiles. He then demonstrated that printing sentences in phrase segments allowed difference-poor readers to reach the comprehension levels of good readers. Presumably segmenting and clustering the sentences into spacedmeaningful chunks compensated for the difference-poor readers' inability to do so, since these students apparently tend to read word by word. For deficit-poor readers, printing sentences as vertical word lists increased comprehension, though their comprehension was still well below that of good readers. Presumably here, the word lists overcame the deficit-poor readers' tendency to skip words they wereunsure of by focusing attention on each word in turn. It should be noted that fragmenting sentences randomly or printing them as word lists disrupted good readers' processing. They did best with conventional sentences. Here may be an example of segmentation, a type of mathemagenic event originally hypothesized by Rothkopf (1965). Note, however, that different kinds of stimulus segmentation are needed for students with different ability profiles.

Short Term Memory. Estes (1974) advocated detailed experimental analyses of the tasks typically found in intelligence tests, and reviewed past research on several examples of this. He listed various sources of individual differences on digit span tasks, such as familiarity with ordinal numbers, grouping strategies, coding processes, and capacity for selective inhibition. On the digit-symbol substitution part of the WAIS, he noted prior research indicating the importance of distinctive verbal coding in memory. Vocabulary test performance was hypothesized to depend on the structure of long term memory built through verbal experience, but also on accessibility and retrieval strategies, and on the individual's notion of what constitutes an acceptable answer. Word naming tests were likened to free recall processes, suggesting the importance of organization and category usage in verbal recall. Process theories of such tasks could lead to process theories of G, and to improved measures for diagnosis and remediation.

Other recent studies relate to some of these suggestions. Mason, Katz, and Wicklund (1975) contrasted memory for order and memory for items in a modified letter span task. Order memory appeared to be more highly related to reading ability (G_c?) than item memory. And, Jensen and Figueroa (1975) performed several studies relating to the Jensen Level I-Level II ability distinction. As predicted, backward digit span scores were more highly related to WISC-IQ (G_c,G_f, and G) than were forward digit span scores. Differences in these relations were stronger for black than for white subjects. The interpretation suggests that backward digit span requires more mental processing steps than does forward digit span and, hence, is closer to Level II transformation ability. Note also its appearance with G_c tests in the Das-Molloy factor study (Table 3).

A program of studies by E. Hunt (See, e.g., Hunt, Frost, & Lunneborg, 1973) has sought to connect G_c to various parameters representing speed and sequence of short-term memory processing. In a visual matching task, High G_c subjects were found to be faster at name matching, but not at physical identity matching, than subjects with lower G_c scores. Another experiment used a continuous paired-associates task, suggesting that G_c related to parameters defined as showing probability of entry of an item into short term memory, and negative rate of loss of items from an intermediate memory storage. In two other studies, High G_c subjects showed more release from proactive inhibition in word recall and shallower slope parameters (i.e., faster memory search) in a Sternberg memory search task, than did subjects with lower G_c scores. Hunt hypothesizes that verbal ability is associated with speed of coding, order-preserving, and search operations in short-term memory, but his ability measures cannot be satisfactorily interpreted as verbal specifically; they represent G_c (See Snow, 1976.).

Chiang and Atkinson (1976, in press) pursued Hunt's findings on the visual and memory search tasks. They failed to demonstrate significant overall correlations between SAT scores (G_c) and the slope parameters indicating speed of search. However, striking relations appeared when the sample of subjects was separated by sex. Higher G_c scores were associated with faster memory search for males (consistent with Hunt), while for females the opposite relation obtained; here however the scatterplot contained some strange outliers.

Mediation and transformation. Studies by Keislar and Stern (1970) compared single hypothesis vs. multiple hypothesis strategies in concept attainment. High G children performed better using multiple hypotheses but Low G children did best with the single hypothesis method.

Cason (1972) has succeeded in developing measures of process individual differences in concept attainment for hypothesis generation, hypothesis evaluation, and memory for hypotheses. But these have not yet been related to mental ability measures or shown to be relevant to experimental manipulations.

A line of studies by Rohwer has investigated various verbal and pictorial elaborations in associative learning, and their relation to age, race, and SES variables. It is thought, for example, that age effects are due to older subjects using more, or more highly developed, mental elaborations (such as sentences) to connect members of the pairs to be learned. Age × treatment interactions have been shown, as well as marked individual differences in elaborative activity (Rohwer and Bean, 1973). In more recent studies (Rohwer, et al., n.d.), both G_c and a specific measure of individual differences in paired-associate learning have been studied in relation to the elaboration hypotheses. It was reported that elaborative propensity varies reliably as an individual difference in adolescents, can be prompted by treatments providing different kinds of elaborative instructions, and seems only weakly related to G_c. There was an implication in one experiment that repetition instructions (designed to impede elaboration) were particularly harmful for subjects with high pretest scores on paired-associate learning proficiency; this may be another case of the interference phenomenon noted in a previous section of this chapter.

Imagery seems to aid associative learning as another form of elaboration. How imagery helps has been the subject of research by Paivio (1974), DiVesta, et al., (1971), and many others. Paivio interprets the data as suggesting that imagery and verbal functions represent a dual processing system. There is evidence that individuals differ in the use and efficiency of the two modes of mediation. In much of this work, however, subjects who are High vs. Low in imagery have been identified using spatial ability tests. Image-generation is probably one of the processes wrapped up in G_v. Kerst and Levin (1973) have looked at individual differences here, showing that imagery and sentence-mediators, both experimenter-given and learner-generated, are equally facilitative of retention, but that learner-generated mediators display markedly greater individual differences. Levin (1973) also gave stories in print vs. pictorial form to readers who had been classified as good readers, difference-poor readers (weak

in comprehension), or deficit-poor readers (weak in both vocabulary and comprehension), following a distinction made by Cromer citer earlier. Half of the print group were also given imagery instructions. Reading with imagery facilitated the comprehension of difference-poor readers, and also of good readers, but not of deficit-poor readers. The latter group did best with pictures.

Part of the research on imagery vs. verbal elaborations stemmed from early findings that pictures were better than words as stimuli in paired-associate learning. Rohwer and Harris (1975) have gone on to suggest that combinations of media (e.g., pictorial augmentation of oral prose) might be particularly beneficial for members of some populations (low SES black children), but not necessarily for other populations. Meanwhile, Levin, et al., (1974) developed a paired associate learning test to identify subjects who learned relatively well from pictures vs. words. It was then shown that learners who were good with pictures, but poor with words, did significantly better when given picture pairs instead of word pairs. A second experiment used a prose reading task given under regular vs. imagery instructions, to compare comprehension of the passage by good picture-learners and poor picture-learners. While no differences among the groups occurred under regular reading instructions, the former subjects were helped, and the latter apparently impeded, by instructions to use imagery.

Coffing (1971) replicated one early experiment by Rohwer (1967) that had compared pictures and words, adding measurement of individual differences in eye movements. His data suggested that pictures facilitate learning only for subjects who tend to scan pictures rather than words when given a choice. Those who seem to prefer to look at the words showed no differences in learning outcome between conditions. Also of note is a study by Hall (1974) suggesting that eye movement patterns relate to the subject's use of iconic imagery. During recall of letters, eye movements tended to scan the spatial positions where the letters had previously appeared. These studies tie back in with those mentioned in the initial stimulus processing section. Individual differences in eye movements may deserve attention as one important source of clues about processing activities in several stages of learning.

Perhaps the simplest example of an experiment designed to examine ATI in the laboratory comes from the work of Gavurin (1967). This and a follow-up study by Frandsen and Holder (1969) suggest that G_v reflects in part skill in mentally transforming and thus solving verbal problems. Gavurin gave some subjects anagram problems under standard conditions and allowed other subjects to work the problems by moving the printed letters about on the table top to try out alternative arrangements. A spatial test correlated strongly with problem solving under the

nonmanipulation condition and essentially zero when subjects were allowed to manipulate the letters. Prohibiting manipulation apparently forces reliance on a kind of internal processing at which High G_v subjects excel.

Frandsen and Holder followed by demonstrating strong correlation between G_v and verbal problem solving, using syllogisms, time-rate-distance and deduction problems, even when G_c was partialled out. They then showed that training in use of diagrammatic techniques like Venn diagrams and time lines brought the problem solving performance of Low G_v subjects up to that of High G_v subjects.

Shephard and his coworkers (see, e.g., Shepard and Feng 1972) have analyzed the processes involved in many tasks comparable to those used to measure G_v. Their work has not been concerned with individual differences, but implies that G_v represents differences in the construction and transformation of mental analogs, and speed therein. While Shepard's work thus seems to argue against a sequential feature comparison account of G_v, French (1965) has shown that subjects divide between these two strategies on the basis of self-report, at least on some spatial tasks. As noted earlier then, the hypothesis is that some subjects solve spatial rotation problems by visual analog transformations (G_v); others solve them by a process of logical-sequential matching of stimulus features (G_f or G_c).

Contrasts between G_v and G_c, or between imagery and verbal processing systems, prompt hypotheses that relate these distinctions to cerebral hemispheric structure. The Das conjectures about simultaneous vs. successive processing, or the Pask-Scott distinction between holist and serialist processing, both noted earlier, are similarly suggestive. Evidence to date suggests that the left cerebral hemisphere concerns verbal-analytic processing, while spatial image processing may be more of a right hemisphere function. A review of research on cerebral hemispheric assymetry by Lohman (1976) shows reliable individual differences in lateralization of language processes, with average differences most notable between sinistrals and dextrals, males and females, and children and adults. It remains to be seen how far neurophysiological measures can be pushed to help clarify psychological process conceptions of such individual differences.

Reasoning and problem-solving. In a rare, early analysis, Simon and Kotovsky (1963; see also Kotovsky and Simon, 1973) produced a computer simulation of performance on the Thurstone Letter Series Test. Their model distinguished two processes, pattern analysis and sequence generation, but dealt with individual differences only indirectly, in a few variants on the basic model. These deserve research attention, however. Other investigators have begun similarly detailed analyses of other tests usually considered fundamental to the definition of G (or G_f).

Whitely (1976) studied a large pool of verbal analogy test items, dividing them by a cluster analysis procedure into eight types of relational concepts. She then administered these, along with a measure of eduction of multiple relationships for noun pairs and six other ability tests, to college students. Predicting ability test scores from multiple regression equations made up of analogy item types suggested that different categories of analogy items tap different abilities. The eduction of relationships measure also showed this pattern. But relations between measures were generally low and the correlational analysis was incomplete.

R. Sternberg (1975) has conducted the most comprehensive review and analysis of a mental test performance yet reported. His empirical work dealt with analogies, but the report includes a valuable review of differential and information processing theory and research relevant to a theory of G (or G_f). Sternberg also developed his own "componential" theory of analogical reasoning, with a detailed task analysis to identify specific components of this form of reasoning, their combination rules, and forms of processing. All this material cannot be synopsized here. While portions are discussed by Snow (1976) in relation to other theoretical and methodological issues the reader is urged to examine the original report in detail.

In brief, Sternberg's analysis produces separate reaction time measures for six components: encoding, inference, mapping, application, justification, and preparation-response. Various alternative models can be used to combine these. A series of three experiments presented pictorial, verbal, and geometric analogies to extreme groups on G_f and perceptual speed, along with other G_c and analogy reference tests. Results suggested that: all subjects use all components except justification, which is included by some subjects but not others; individual

differences arise from differing component times, not from different combination rules or models; G_f sometimes correlates positively with encoding time, suggesting a strategy difference whereby more time taken by Highs in encoding makes later component operation more efficient; G_f correlates negatively with time for each later component, but only under specified conditions, as e.g., when inference and mapping involve discovery of relevant attributes rather than just their testing; perceptual speed sometimes correlates negatively with encoding time, but this reference ability is generally not associated with components of the theory.

A few other studies can be inserted here for future reference. Jacobs and Vandeventer (1971a; 1971b), Guinaugh (1969), and Turner (1975) have all conducted training experiments on Raven Progressive Matrices (G_f). These should help guide process analyses of this task of the sort that Sternberg applied to analogies.

Strategies and structure. Research by Dunham and Bunderson (1969) connects to some extent with the Greeno-Mayer work reviewed earlier. It also introduces research distinguishing G from memory abilities and raises the large issue of compensatory plans or strategies for using one's abilities. One of their investigations showed that a principle-structured treatment depended particularly on reasoning ability (G_f?), while a general orientation or relatively unstructured treatment depended on memory ability, using simple concept-attainment tasks. The implication is that what ability determines success depends on the way one stores incoming information. If one must remember isolated bits of information, memory ability is the necessary condition for learning. If one can sort the relevant information by applying some rational structure or principle then only this information needs to be remembered. Memory load is reduced but now the ability to reason is the necessary condition. Each treatment serves a compensatory function; each allows the learner to substitute an ability he has for one he lacks.

Dunham (1969) reported a similar finding comparing a condition where pairs of stimuli presented together were always instances of the same concept, with a condition where pairs were instances of different concepts. Reasoning ability correlated with performance in the "same" condition while memory ability correlated with performance in the "different" condition.

Merrill (1974) added the finding that a negative relation of G_f to test item response latency, found in an examples-only condition, was reduced under conditions where objectives and/or rules were provided. The task was brief CAI instruction about an imaginary science.

Labouvie, et al (1973) used an immediate recall vs. delayed recall contrast with picture stimuli. A series of 30 pictures were shown on each trial, with a letter cancellation task performed in the delay interval. Subjects were college students. Memory ability measures correlated with recall performance primarily on early blocks of trials in the immediate recall condition. G measures correlated with performance mainly on late trial blocks in the delay condition. In a later series of experiments, Labouvie-Vief, et al (1975) tried to replicate this finding. One study gave a similar recall task under three conditions: standard free recall, chunking instructions, rote rehearsal instructions. Here subjects were from Grades 7 and 12. G correlated with recall performance on all trials in all conditions at both age levels. Memory measures correlated with nothing. Another study used subjects from Grades 3 and 10, using a recall task and delay vs. no delay conditions like that in the first study. Raven (G_f) and digit span were the only ability measures (whereas earlier work had included multiple measures and factor scores for the aptitude constructs). Results here showed scattered significant correlations of recall with G_f but none with memory span. Finally, a year later the third graders above performed a paired associate task. This was given under conditions designed to promote speeded rote memory vs. an imagery strategy in learning the pairs. G_f correlated significantly with performance in the conditions prompting use of an imagery strategy and zero in the speeded rote condition. Digit span gave moderate nonsignificant correlations in all conditions.

Thus, G or G_f seems relevant to recall performance under various conditions at various ages, perhaps especially when subjects can choose or are prompted to use some organizing strategy. Memory ability showed no consistent relevance.

There have also been studies of individual differences in subjective strategies in word recall. Frederiksen (1969) found that performance in free recall was positively related to G_c and memory span, and negatively related to the learner's self-reported tendency to modify strategies during learning. When subjects had to anticipate words in clusters, those high in fluency and low in semantic flexibility and using a strategy of active sequential organization, did best. In a serial anticipation condition, only use of a mnemonic strategy was related to success.

Cohen (1973) gave paired associates to college students who reported the organizing strategies they used for each item, after studying and recalling several lists. They were then assigned to conditions, designed to promote repetitive (rote), formal (using common letter patterns), or experimental

(using familiar ideas) organizing strategies. A control group was told to use any preferred procedure. It was found that subjects who naturally used more high level strategies (formal or experimental) showed higher recall performance than did subjects preferring repetitive or unconscious strategies. The latter's performance improved under instructions to use high level strategies. Ability measures were not included however.

Scattered other studies have investigated individual differences in subjective organization in recall, but have not been concerned with aptitude constructs. They are difficult to interpret in the context of this chapter. And, Sternberg and Tulving (in press) have shown that different measures of subjective organization in free recall have different psychometric properties, correlate differently with recall, and that these differences account for disparity among the results of a range of experiments.

Response integration and retention. There is little to go on regarding process individual differences in response integration, outside of psychomotor research (For this, see Fleishman, 1975). But Sternberg hypothesized a few additional components in his preparation-response category to account for an unexpected result in his analogy experiments. The component score for preparation-response was highly correlated with G_f, suggesting that preparation and response need to be separated. Then, such a correlation might be found to arise from individual differences in planning at the start of solution, "bookkeeping" during solution, or a decision process occuring near the end of solution. Individual differences in motivation or attention throughout solution were also noted as possible explanations. One can imagine all these processes involved in more general learning and problem solving, but it seems that "response integration" still would deserve more detailed analysis.

Recent studies suggesting process conceptions of individual differences in long-term retention have not been found. There are a few implications to be noted in the Gagné (1967) symposium, and in Cronbach and Snow. And some recent work on short-term retention suggests that learners who differ in their rates of initial learning nonetheless forget at the same rate (Shuell and Giglio, 1973). If there are individual differences in long term retention, they should relate to G_c, since knowledge accumulation and use over long intervals is part of the definition of that construct.

Prospectus

Instructional studies of ATI, methodological improvements for such studies, and process analyses of aptitude, constitute a three-fold path toward the integration of experimental and differential psychology in educational research and, hopefully, toward improved theory and practice to explain and use aptitude for instruction. The first aspect of this path attacks ATI phenomena in school learning directly, the second develops the means to understand ATI data at this level, and the third seeks to transmit to such research improved process models and measures of aptitude. Continuing work relevant to each aspect will likely deserve its own, more intensive review in the coming years. A detailed summary of this chapter here thus seems unnecessary as well as impossible.

In the past, individual investigators have usually concentrated on school studies, or on laboratory work, or on methodology, as <u>their</u> aptitudes have moved them. The three-fold path conception of this chapter is meant to suggest that all three aspects need to be addressed in the future, in coordinated research programs on common ATI phenomena.

Educational researchers interested in this problem can find many different hypotheses on which to focus their work. This chapter has suggested that the $A_iA_cA_x$ and $G_cG_fG_v$ aptitude complexes, and their interconnections, deserve particular attention. Educational researchers not interested in this problem can ignore it. But individual differences in aptitude for learning will not go away.

References

Aiken, L. R., Jr. Verbal factors and mathematics learning: A review of research. Journal for Research in Mathematics Education, 1971, 2, 304-313.

Amick, D. J., & Walberg, H. J. Introductory multivariate analysis. Berkeley, Ca.: McCutchan, 1975.

Anderson, G. L. A comparison of the outcomes of instruction under two theories of learning. Unpublished doctoral dissertation, University of Minnesota, 1941.

Berliner, D. C., & Cahen, L. S. Trait-treatment interaction and learning. In F. N. Kerlinger (Ed.), Review of research in education (Vol. 1). Itasca, Ill.: Peacock, 1973.

Boersma, F. J., Muir, W., Wilton, K., & Barham, R. Eye movements during embedded figure tasks. Perceptual and Motor Skills, 1969, 28, 271-274.

Borich, G., Godbout, R., & Wunderlich, K. The analysis of aptitude-treatment interactions: Computer programs and calculations. Chicago, Ill.: International Educational Services, 1976.

Brownell, W. A., & Moser, A. G. Meaningful versus mechanical learning: A study in grade three subtraction. Duke University Research Studies in Education (No. 8). Durham, N.C.: Duke University Press, 1949.

Campbell, D. T., & Fiske, D. W. Convergent and discriminant validation by the multitrait-multimethod matrix. Psychological Bulletin, 1959, 56, 81-105.

Cason, G. J. Individual differences in preference, saliency, and concept identification performance. Paper presented at the meeting of the American Educational Research Association, Chicago, 1972.

Cattell, R. B. Abilities: Their structure, growth, and action. Boston: Houghton-Mifflin, 1971.

Chiang, A., & Atkinson, R. C. Individual differences and interrelationships among a select set of cognitive skills. Memory and Cognition, in press.

Coffing, D. G. Eye movement preferences as individual differences in learning. Unpublished doctoral dissertation, Stanford University, 1971.

Cohen, S. R. Influence of organizing strategies and instructions on short-term retention. Journal of Educational Psychology, 1973, 64, 199-205.

Conklin, R. C., Muir, W., & Boersma, F. J. Field dependency-independency and eye-movement patterns. Perceptual and Motor Skills, 1968, 26, 59-65.

Crist-Whitzel, J. L, & Hawley-Winne, B. J. Individual differences and mathematics achievement: An investigation of aptitude-treatment interactions in an evaluation of three instructional approaches. Paper presented at the meeting of the American Educational Research Association, San Francisco, April 1976.

Cromer, W. The difference model: A new explanation for some reading difficulties. Journal of Educational Psychology, 1970, 61, 471-483.

Cronbach, L. J. The two disciplines of scientific psychology. American Psychologist, 1957, 12, 671-684.

Cronbach, L. J. Beyond the two disciplines of scientific psychology. American Psychologist, 1975, 30, 116-127.

Cronbach, L. J. Research on classrooms and schools: Formulation of questions, design, and analysis. Stanford University School of Education, Stanford Evaluation Consortium, 1976.

Cronbach, L. J., & Snow, R. E. Aptitudes and instructional methods: A handbook for research on interactions. New York: Irvington, in press.

Cronbach, L. J., & Webb, N. Between-class and within-class effects in a reported aptitude x treatment interaction: Reanalysis of a study by G. L. Anderson. Journal of Educational Psychology, 1975, 67, 717-724.

Das, J. P., & Malloy, G. N. Varieties of simultaneous and successive processing in children. Journal of Educational Psychology, 1975, 67, 213-220.

Das, J. P., Kirby, J., & Jarman, R. F. Simultaneous and successive syntheses: An alternative model for cognitive abilities. Psychological Bulletin, 1975, 82, 87-103.

Day, R. S. Temporal order judgments in speech: Are individuals language-bound or stimulus-bound? Paper presented at the Ninth Annual Meeting of the Psychonomic Society, St. Louis, 1969.

Day, R. S. Digit-span memory in language-bound and stimulus-bound subjects. Journal of the Acoustical Society of America, 1973, 54, 287. (Abstract) Also in Haskins Laboratories Status Report (SR-34), 1973. (a)

Day, R. S. On learning "secret languages." Haskins Laboratories Status Report (SR-34), 1973. (b)

Di Vesta, F. J., Sanders, N. M., Schultz, C. B., & Weener, P. D. Instructional strategies: Multivariable studies of psychological processes related to instruction. Annual report, Part II. Unpublished report. Pennsylvania State University, 1971.

Domino, G. Differential predictions of academic achievement in conforming and independent settings. Journal of Educational Psychology, 1968, 59, 256-260.

Domino, G. Interactive effects of achievement orientation and teaching style on academic achievement. Journal of Educational Psychology, 1971, 62, 427-431.

Domino, G. Aptitude by treatment interaction effects in college instruction. Paper presented at the meeting on the American Psychological Association New Orleans, La., August 30-September 3, 1974.

Dowaliby, F. J., & Schumer, H. Teacher-centered vs. student-centered mode of college classroom instruction as related to manifest anxiety. Journal of Educational Psychology, 1973, 64, 125-132.

Dunham, J. L. Investigations of the role of intellectual abilities in concept learning. Paper presented at the meeting of the American Educational Research Association, Los Angeles, February 1969.

Dunham, J. L., & Bunderson, C. V. The effect of decision-rule instruction upon the relationship of cognitive abilities to performance in multiple-category concept problems. Journal of Educational Psychology, 1969, 60, 121-125.

Edgerton, H. A. Should theory precede or follow a "How-to-do-it" phase of training? Unpublished report, Richardson, Bellows, Henry, & Co., New York, 1956.

Edgerton, H. A. The relationship of method of instruction to trainee aptitude pattern. Unpublished report, Richardson, Bellows, Henry, & Co., New York, 1958.

Egan D. E., & Greeno, J. G. Acquiring cognitive structure by discovery and rule learning. Journal of Educational Psychology, 1973, 64, 85-97.

Estes, W. K. Learning theory and intelligence. American Psychologist, 1974, 29, 740-749.

Frandsen, A. N., & Holder, J. R. Spatial visualization in solving complex problems. Journal of Psychology, 1969, 73, 229-233.

Frederiksen, C. H. Abilities, transfer and information retrieval in verbal learning. Multivariate Behavioral Research Monographs, No. 2, 1969.

French, J. W. The relationship of problem-solving styles to the factor composition of tests. Educational and Psychological Measurement, 1965, 25, 9-28.

Fleishman, E. A. Toward a taxonomy of human performance. American Psychologist, 1975, 30, 1127-1149.

Gagné, R. M. (Ed.), Learning and individual differences. Columbus, Ohio: Merrill, 1967.

Gaudry, E., & Spielberger, C. D. Anxiety and educational achievement. New York: Wiley, 1971.

Gavurin, E. I. Anagram solving and spatial aptitude. Journal of Psychology, 1967, 65, 65-68.

Glaser, R. Some implications of previous work on learning and individual differences. In R. M. Gagné (Ed.), Learning and individual differences. Columbus, Ohio: Merrill, 1967.

Glaser. R. Individuals and learning: The new aptitudes. Educational Researcher, 1972, 1, 5-12.

Glaser, R. Components of a psychology of instruction: Toward a science of design. Review of Educational Research, 1976, 46, 1-24.

Greene, J. Choice behavior and its consequences for learning: An ATI study. Unpublished doctoral dissertation, Stanford University, 1976.

Greeno, J. G., & Mayer, R. E. Structural and quantitative interaction among aptitudes and instructional treatments. Unpublished paper, University of Michigan, 1975.

Guilford, J. P. The nature of human intelligence. New York: McGraw-Hill, 1967.

Guinaugh, B. J. An experimental study of basic learning ability and intelligence in low socio-economic populations. Unpublished doctoral dissertation, Michigan State University, 1969.

Gustafsson, J.-E. Verbal versus figural in aptitude-treatment interactions. Review of the literature and an empirical study. Reports from the Institute of Education, University of Göteborg, No. 36, 1974.

Gustafsson, J. -E. Verbal and figural aptitudes in relation to instructional methods: Studies in aptitude-treatment interactions. Göteborg Studies in Educational Sciences 17, Göteborg, Sweden, 1976.

Guttman, L. The first law of intelligence and attitude. Paper presented at the meeting of the American Educational Research Association, San Francisco, April 1976.

Hall, D. C. Eye movements in scanning iconic imagery. Journal of Experimental Psychology, 1974, 103, 825-830.

Helson, H. Adaptation-level theory. New York: Harper and Row, 1964.

Hettema, J. Cognitive abilities as process variables. Research Bulletin No. RB-67-6 Princeton, N.J.: Educational Testing Service, 1967.

Horn, J. L. Human abilities: A review of research and theory in the early 1970s. Annual Review of Psychology, 1976, 27, 437-485.

Hunt, D. E. Person-environment interaction: A challenge found wanting before it was tried. Review of Educational Research, 1975, 45, 209-230.

Hunt, E. B., Frost, N., & Lunneborg, C. E. Individual differences in cognition: A new approach to intelligence. In G. Bower (Ed.), Advances in learning and memory (Vol. 7). New York: Academic Press, 1973.

Jacobs, P. I., & Vandeventer, M. The learning and transfer of double-classification skills by first graders. Child Development, 1971, 42, 149-159. (a)

Jacobs, P. I., & Vandeventer, M. The learning and transfer of double-classification skills: A replication and extension. Journal of Experimental Child Psychology, 1971, 12, 240-257. (b)

Jensen, A. R. Varieties of individual differences in learning. In R. M. Gagné (Ed.), Learning and individual differences. Columbus, Ohio: Merrill, 1967.

Jensen, A. R., & Figueroa, R. A. Forward and backward digit span interaction with race and IQ: Predictions from Jensen's theory. Journal of Educational Psychology, 1975, 67, 882-893.

Keislar, E. R., & Stern, C. Differentiated instruction in problem solving for children of different mental ability levels. Journal of Educational Psychology, 1970, 61, 445-450.

Kerlinger, F. N., & Pedhazur, E. J. Multiple regression in behavioral research. New York: Holt, Rinehart and Winston, 1973.

Kerst, S., & Levin, J. R. A comparison of experimenter-provided and subject-generated strategies in children's paired-associate learning. Journal of Educational Psychology, 1973, 65, 300-303.

Klitgaard, R. E. Looking for the best. Santa Monica, Ca.: The Rand Corporation, 1976.

Kotovsky, K., & Simon, H. A. Empirical tests of a theory of human acquisition of concepts for sequential patterns. Cognitive Psychology, 1973, 4, 399-424.

Labouvie, G. V., Frohring, W. R., Baltes, P. B., & Goulet L. R. Changing relationship between recall performance and abilities as a function of stage of learning and timing of recall. Journal of Educational Psychology, 1973, 64, 191-198.

Labouvie-Vief, G., Levin, J. R., & Urberg, K. A. The relationship between selected cognitive abilities and learning: A second look. Journal of Educational Psychology, 1975, 67, 558-569.

Levin, J. R. Inducing comprehension in poor readers: A test of a recent model. Journal of Educational Psychology, 1973, 65, 19-24.

Levin, J. R., Divine-Hawkins, P., Kerst, S. M., & Guttmann, J. Individual differences in learning from pictures and words: The development and application of an instrument. Journal of Educational Psychology, 1974, 66, 296-303.

Lohman, D. F. Lateral differences in ability: A review. Technical Report No. 4. Aptitude Research Project, School of Education, Stanford University (in preparation).

Marks, M. R. How to build better theories, tests, and therapies: The off-quadrant approach. American Psychologist, 1964, 19, 793-798.

Mason, M., Katz, L., & Wicklund, D. A. Immediate spatial order memory in sixth grade children as a function of reader ability. Journal of Educational Psychology, 1975, 67, 610-616.

Mayer, R. E. Different problem-solving competencies established in learning computer programming with and without meaningful models. Journal of Educational Psychology, 1975, 67, 725-734.

Mayer, R. E., Stiehl, C. C., & Greeno, J. G. Acquisition of understanding and skill in relation to subjects' preparation and meaningfulness of instruction. Journal of Educational Psychology, 1975, 67, 331-350.

McCarthy, S. V. Differential V-Q ability: Twenty years later. Review of Educational Research, 1975, 45, 263-282.

Merrill, P. F. Effects of the availability of objectives and/or rules on the learning process. Journal of Educational Psychology, 1974, 66, 534-539.

Myers, A. E. Risk taking and academic success and their relation to an objective measure of achievement motivation. Research Bulletin No. RB-64-2 Princeton, N.J.: Educational Testing Service, 1964.

Neisser, U. Cognitive psychology. New York: Appleton-Century-Crofts, 1967.

Paivio, A. Language and knowledge of the world. Educational Researcher, 1974, 3, 5-12.

Pask, G., & Scott, B. C. E. Learning strategies and individual competence. International Journal of Man-Machine Studies, 1972, 4, 217-253.

Peterson, P. L. Interactive effects of student anxiety, achievement orientation, and teacher behavior on student achievement and attitude. Unpublished doctoral dissertation, Stanford University, 1976.

Porteus, A. Teacher-centered vs. student-centered instruction: Interactions with cognitive and motivational aptitudes. Unpublished doctoral dissertation, Stanford University, 1976.

Resnick, L. (Ed.), The nature of intelligence. Hillsdale, N.J.: Lawrence Erlbaum Associates, Inc., 1975.

Rohwer, W. D., Jr. Social class differences in the role of linguistic structures in paired-associate learning. Office of Education Contract No. 6-10-273, Berkeley, Ca.: University of California, 1967.

Rohwer, W. D., Jr., & Bean, J. P. Sentence effects and noun-pair learning: A developmental interaction during adolescence. Journal of Experimental Child Psychology, 1973, 15, 521-533.

Rohwer, W. D., Jr., & Harris, W. J. Media effects on prose learning in two populations of children. Journal of Educational Psychology, 1975, 67, 651-657.

Rohwer, W. D., Jr., Raines, J. M., Eoff, J., & Wagner, M. The development of elaborative propensity in adolescence. Unpublished paper, University of California at Berkeley, no date.

Rothkopf, E. Z. Some theoretical and experimental approaches to problems in written instruction. In J. D. Krumboltz (Ed.), Learning and the educational process. Chicago: Rand McNally, 1965.

Rothkopf, E. Z. Symposium discussion: Individual differences in learning related processes. Paper presented at the meeting of the American Educational Research Association, Chicago, April 1972.

Rutkowski, K., & Domino, G. Interrelationship of study skills and personality variables in college students. Journal of Educational Psychology, 1975, 67, 784-789.

Salomon, G. Interaction of communication-medium and two procedures of training for subjective response uncertainty of teachers. Unpublished doctoral dissertation, Stanford University, 1968.

Salomon, G. Internalization of filmic schematic operations in interaction with learners' aptitudes. Journal of Educational Psychology, 1974, 66, 499-511.

Satterly, D. J. Cognitive styles, spatial ability, and school achievement. Journal of Educational Psychology, 1976, 68, 43-47.

Seibert, W. F., & Snow, R. E. Studies in cine-psychometry I: Preliminary factor analysis of visual cognition and memory. Final Report, USOE Grant Number 7-12-0280-184. Lafayette, Ind.: Audio Visual Center, Purdue University, 1965.

Shepard, R. N., & Feng, C. A chronometric study of mental paper folding. Cognitive Psychology, 1972, 3, 228-243.

Shuell, T. J., & Giglio, J. Learning ability and short-term memory. Journal of Educational Psychology, 1973, 64, 261-266.

Simon, H. A., & Kotovsky, K. Human acquisition of concepts for sequential patterns. Psychological Review, 1963, 70, 534- 546.

Skanes, G. R., Sullivan, A. M., Rowe, E. J., & Shannon, E. Intelligence and transfer: Aptitude by treatment interactions. Journal of Educational Psychology, 1974, 66, 563-568.

Snow, R. E. Representative and quasi-representative designs for research on teaching. Review of Educational Research, 1974, 44, 265-292.

Snow, R. E. Theory and method for research on aptitude processes. Technical Report No. 2. Aptitude Research Project, School of Education, Stanford University, 1976.

Snow, R. E., & Seibert, W. F. Exploratory factor analysis in cine-psychometry. Paper presented at the meeting of the American Psychological Association, New York, September 1966.

Spielberger, C. D. Conceptual and methodological issues in anxiety, drive theory, and computer-assisted learning. In B. A. Maher, (Ed.), Progress in experimental personality research (Vol. 6). New York: Academic Press, 1972.

Sternberg, R. J. The componential analysis of human abilities: Intelligence, information processing, and analogical reasoning. Unpublished doctoral dissertation, Stanford University, 1975.

Sternberg, R. J., & Tulving, E. The measurement of subjective organization in free recall. Psychological Bulletin, in press.

Sticht, T. G. Learning by listening in relation to aptitude, reading, and rate-controlled speech: Additional studies. (HumRRO Technical Report 71-5) Alexandria, Va.: Human Resources Research Organization, 1971.

Sullivan, H. J., Okada, M., & Niedermeyer, F. C. Learning and transfer under two methods of word-attack instruction. American Educational Research Journal, 1971, 8, 227-239.

Tobias, S. Achievement treatment interactions. Review of Educational Research, 1976, 46, 61-74.

Trabasso, T., & Bower, G. Attention in learning: Theory and research. New York: Wiley, 1968.

Turner, R. R. Effect of information feedback on matrices learning for differing socio-economic levels. Journal of Educational Psychology, 1975, 67, 285-295.

Whitely, S. E. Solving verbal analogies: Some cognitive components of intelligence test items. Journal of Educational Psychology, 1976, 68, 117-127.

Wijnen, W. H. F. W., & Snow, R. E. Implementing an evaluation system for medical education. Technical Report No. 1. Medische Faculteit Maastricht, Maastricht, Netherlands, 1975.

Distribution List

4 Dr. Marshall J. Farr, Director
Personnel & Training
Research Programs
Office of Naval Research
Arlington, VA 22217

1 Director
ONR Branch Office
495 Summer Street
Boston, MA 02210
ATTN: Dr. James Lester

1 Director
ONR Branch Office
1030 East Green Street
Pasadena, CA 91101
ATTN: E. E. Gloye

1 Director
ONR Branch Office
536 South Clark Street
Chicago, IL 60605
ATTN: Dr. Charles Davis

1 Dr. M. A. Bertin
Scientific Director
Office of Naval Research
Scientific Liaison Group/Tokyo
APO San Francisco 96503

1 Office of Naval Research
Code 200
Arlington, VA 22217

1 Assistant Deputy Chief of Naval
Personnel for Retention Analysis
and Coordination (Pers 12)
Room 2403, Arlington Annex
Washington, DC 20370

1 Dr. Lee Miller
Naval Air Systems Command
AIR-413E
Washington, DC 20361

1 Commanding Officer
Naval Health Research Center
San Diego, CA 92152
ATTN: Library

1 Director, Navy Occupational Task
Analysis Program (NOTAP)
Navy Personnel Program Support
Activity
Building 1304, Bolling AFB
Washington, DC 20336

1 Office of Civilian Manpower Management
Code 263
Washington, DC 20390

1 Chief of Naval Reserve
Code 3055
New Orleans, LA 70146

1 Chief of Naval Operations
OP-987P7
Washington, DC 20350
ATTN: Capt. H. J. M. Connery

1 Director
Training Analysis & Evaluation Group
Code N-00t
Department of the Navy
Orlando, FL 32813
ATTN: Dr. Alfred F. Smode

1 LCDR C. F. Logan, USN
F-14 Management System
COMFITAEWWINGPAC
NAS Miramar, CA 92145

5 Navy Personnel Research
and Development Center
Code 01
San Diego, CA 92152

1 Navy Personnel Research
and Development Center
Code 02
San Diego, CA 92152
ATTN: A. A. Sjoholm

1 Navy Personnel Research
and Development Center
Code 306
San Diego, CA 92152
ATTN: Dr. J. McGrath

1 Navy Personnel Research
and Development Center
San Diego, CA 92152
ATTN: Library

1 Navy Personnel Research
and Development Center
Code 9041
San Diego, CA 92152
ATTN: Dr. J. D. Fletcher

1 D. M. Gragg, CAPT, MC, USN
Head, Educational Programs
Development Department
Naval Health Sciences Education
and Training Command
Bethesda, MD 20014

5 Director
Naval Research Laboratory
Code 2627
Washington, DC 20390

12 Defence Documentation Center
Cameron Station, Building 5
5010 Duke Street
Alexandria, VA 22314

1 Chairman
Behavioral Science Department
Naval Command and Management Div.
U.S. Naval Academy
Luce Hall
Annapolis, MD 21402

1 Chief of Naval Technical Training
Naval Air Station Memphis (75)
Millington, TN 38054
ATTN: Dr. Norman J. Kerr

1 Chief of Naval Training
Naval Air Station
Pensacola, FL 32508
ATTN: Dr. W. Scanland

1 Command Officer
U.S. Naval Amphibious School
Coronado, CA 92155

1 Dr. James J. Regan
Technical Director
Navy Personnel Research
and Development Center
San Diego, CA 92152

1 Chief
Bureau of Medicine and Surgery
Code 413
Washington, DC 20372

1 Mr. Arnold Rubinstein
Naval Material Command (0344)
Room 1044, Crystal Plaza 5
Washington, DC 20360

1 Superintendent
Naval Postgraduate School
Monterey, CA 93940
ATTN: Library (Code 2124)

1 Chief of Naval Training Support
Code N-21, Building 45
Naval Air Station
Pensacola, FL 32508

1 Armed Forces Staff College
Norfolk, VA 23511
ATTN: Library

1 Mr. George N. Graine
Naval Ship Systems Command
SEA 047C12
Department of the Navy
Washington, DC 20362

1 Commanding Officer
Service School Command
U.S. Naval Training Center
San Diego, CA 92133
ATTN: Code 3030

1 Dr. William L. Maloy
Principal Civilian Advisor
for Education and Training
Naval Training Command, Code OOA
Pensacola, FL 32508

1 HQ USAREUR & 7th Army ODCSOPS
USAREUR Director of GED
APO New York 09403

1 ARI Field Unit - Leavenworth
Post Office Box 3122
Fort Leavenworth, KS 66027

1 Dr. Milton S. Katz, Chief
Individual Training and
Performance Evaluation
U.S. Army Research Institute for
the Behavioral Social Sciences
1300 Wilson Boulevard
Arlington, VA 22209

1 Technical Director
U.S. Army Research
Institute for the Behavioral
and Social Sciences
1300 Wilson Boulevard
Arlington, VA 22209

1 Dr. Frank Harris
U.S. Army Research Institute for
the Behavioral and Social Sciences
1300 Wilson Boulevard
Arlington, VA 22209

1 Dr. Joseph Ward
U.S. Army Research Institute for
the Behavioral and Social Sciences
1300 Wilson Boulevard
Arlington, VA 22209

1 Mr. James Baker
U.S. Army Research Institute for
the Behavioral and Social Sciences
1300 Wilson Boulevard
Arlington, VA 22209

1 Commandant
United States Army Infantry School
ATTN: ATSH - DET
Fort Benning, GA 31905

1 U.S. Army Research Institute
Commonwealth Building - Room 239
1300 Wilson Boulevard
Arlington, VA 22209
ATTN: Dr. R. Dusek

1 Dr. Leon H. Nawrocki
U.S.Army Research Institute
Rosslyn Commonwealth Building
1300 Wilson Boulevard
Arlington, VA 22209

1 Dr. Marty Rockway
Technical Training Division
Lowry Air Force Base
Denver, CO 80230

1 Research Branch
AF/DPMYAR
Randolph AFB, TX 78148

1 AFHRL/DOJN
Stop 63
Lackland AFB, TX 78236

1 Dr. Alfred R. Fregly
Air Force Office of Scientific
Research/PM
Bolling Air Force Base
Washington, DC 20032

1 Capt. Jack Thorpe, USAF
Flying Training Division
AFHRL/FT
Williams AFB, AZ 85224

1 AFHRL/PED
Stop 63
Lackland AFB, TX 78236

1 Instructional Technology Branch
AF Human Resources Laboratory
Lowry AFB, CO 80230

1 AFHRL/AS (Dr. G. A. Eckstrand)
Wright-Patterson Air Force Base
OH 45433

1 AFHRL (AST/Dr. Ross L. Morgan)
Wright-Patterson Air Force Base
OH 45433

1 Headquarters Electronic Systems Div.
ATTN: Dr. Sylvia R. Mayer/MCIT
LG Hanscom Field
Bedford, MA 01730

1 Director, Office of Manpower
Utilization
HQ, Marine Corps (Code MPU)
MCB (Building 2009)
Quantico, VA 22134

1 Chief, Academic Department
Education Center
Marine Corps Development
and Education Command
Marine Corps Base
Quantico, VA 22134

1 Mr. E. A. Dover
2711 South Veitch Street
Arlington, VA 22206

1 Dr. A. K. Slafkosky
Scientific Advisor (Code RD-1)
Commandant of the Marine Corps
Washington, DC 20380

1 Mr. Joseph J. Cowan, Chief
Psychological Research Branch (P-1)
U.S. Coast Guard Headquarters
400 Seventh Street, SW
Washington, DC 20590

1 Dr. John Ford, Jr.
Navy Personnel Research
and Development Center Code 304
San Diego, CA 92152

1 LCDR Charles Theisen, Jr.,
MSC, USN, 4024
Naval Air Development Center
Warminster, PA 18974

1 Military Assistant for Human
Resources
Office of the Secretary of Defense
Room 3D129, Pentagon
Washington, DC 20301

1 Advanced Research Projects Agency
Administrative Services
1400 Wilson Boulevard
Arlington, VA 22209
ATTN: Ardella Holloway

1 Dr. Harold F. O'Neil, Jr.
Advanced Research Projects Agency
Cybernetics Technology Office
1400 Wilson Boulevard
Arlington, VA 22209

1 Dr. Robert Young
Advanced Research Projects Agency
Cybernetics Technology Office
1400 Wilson Boulevard
Arlington, VA 22209

1 Dr. William Gorham, Director
Personnel Research & Development Ctr
U.S. Civil Service Commission
1900 E Street, N.W.
Washington, DC 20415

1 Dr. Marshall S. Smith
Assistant Acting Director
Program on Essential Skills
National Institute of Education
Brown Building - Room 815
19th and M Streets, N.W.
Washington, DC 20208

1 Dr. Carl Frederiksen
Learning Division, Basic Skills Grp
National Institute of Education
1200 - 19th Street, N.W.
Washington, DC 20208

1 Dr. Eric McWilliams
Program Manager
Technology and Systems, TIE
National Science Foundation
Washington, DC 20550

1 Dr. Scarvia Anderson
Executive Director for Special
Development
Educational Testing Service
Princeton, NJ 08540

1 Professor Keith Wescourt
Stanford University
Inst. for Mathematical Studies
in the Social Sciences
Stanford, CA 94305

1 Century Research Corporation
4113 Lee Highway
Arlington, VA 22207

1 Dr. Robert Glaser, Director
University of Pittsburgh
Learning Research and Development Ctr
Pittsburgh, PA 15213

1 Dr. Kenneth E. Clark
University of Rochester
College of Arts and Sciences
River Campus Station
Rochester, NY 14627

1 ERIC
Processing and Reference Facility
4833 Rugby Avenue
Bethesda, MD 20014

1 Dr. Victor Fields
Department of Psychology
Montgomery College
Rockville, MD 20850

1 Dr. Henry J. Hamburger
University of California
School of Social Sciences
Irvine, CA 92664

1 Dr. M. D. Havron
Human Sciences Research, Inc.
Westgate Industrial Park
7710 Old Springhouse Road
McLean, VA 22101

1 Dr. Earl Hunt
Department of Psychology
University of Washington
Seattle, WA 98195

1 Dr. Lawrence B. Johnson
Lawrence Johnson and Assoc., Inc.
200 S. Street, NW, Suite 502
Washington, DC 20009

1 Dr. David Klahr
Carnegie-Mellon University
Graduate School of Industrial Adm.
Pittsburgh, PA 15213

1 Dr. Robert R. Mackie
Human Factors Research Inc.
6780 Cortona Drive
Santa Barbara Research Park
Goleta, CA 93017

1 Dr. Andrew R. Molnar
Technological Innovations in Educ.
National Science Foundation
Washington, DC 20550

1 Dr. Leo Munday
Vice President
Americal College Testing Program
P. O. Box 168
Iowa City, IA 52240

1 Dr. Donald A. Norman
University of California, San Diego
Center for Human Information
Processing
La Jolla, CA 92037

1 Mr. Luigi Petrullo
2431 North Edgewood Street
Arlington, VA 22207

1 Dr. Diane M. Ramsey-Klee
R-K Research & System Design
3947 Ridgemont Drive
Malibu, CA 90265

1 Dr. Joseph W. Rigney
Behavioral Technology Laboratories
University of Southern California
3717 South Grand
Los Angeles, CA 90007

1 Dr. Andrew M. Rose
American Institutes for Research
Suite 200
1055 Thomas Jefferson Street, NW
Washington, DC 20007

1 Dr. Leonard L. Rosenbaum,
Chairman
Department of Psychology
Montgomery College
Rockville, MD 20850

1 Dr. George E. Rowland
Rowland and Company, Inc.
P. O. Box 61
Haddonfield, NJ 08033

1 Dr. Arthur I. Siegel
Applied Psychological Services
and Science Center
404 East Lancaster Avenue
Wayne, PA 19087

1 Dr. Benton J. Underwood
Northwestern University
Department of Psychology
Evanston, IL 60201

1 Dr. David J. Weiss
University of Minnesota
Depatment of Psychology
Minneapolis, MN 55455

1 Dr. Anita West
Denver Reserach Institute
University of Denver
Denver, CO 80210

1 Dr. Richard Atkinson
National Science Foundation
Washington, DC 20550

1 Dr. William Mann
Information Sciences Institute
4676 Admiralty Way
Marina del Ray, CA 90291

1 Dr. John Annett
Department of Psychology
The University of Warwick
Coventry CV47AL
ENGLAND

1 Mr. Samuel Ball
Educational Testing Service
Princeton, NJ 08540

1 Dr. Gerald V. Barrett
University of Akron
Department of Psychology
Akron, OH 44325

1 Dr. Ronald P. Carver
School of Education
University of Missouri-Kansas City
5100 Rockhill Road
Kansas City, MO 64110

1 Dr. Ruth Day
Yale University
Department of Psychology
2 Hillhouse Avenue
New Haven, CT 06520

1 Dr. Edwin A. Fleishman
American Institutes for Research
Suite 200
1055 Thomas Jefferson Street, NW
Washington, DC 20007

1 HumRRO/Central Division
400 Plaza Building
Pace Boulevard at Fairfield Drive
Pensacola, FL 32505

1 HumRRO/Western Division
27857 Berwick Drive
Carmel, CA 93921
ATTN: Library

1 HumRRO/Central Division
Columbus Office
Suite 23
2601 Cross Country Drive
Columbus, GA 31906

1 HumRRO/Western Division
27857 Berwick Drive
CArmel, CA 93921
ATTN: Dr. Robert Vineberg

1 Dr. Arnold F. Kanarick
Honeywell, Inc.
2600 Ridge Parkway
Minneapolis, MN 55413

1 Dr. Steven W. Keele
University of Oregon
Department of Psychology
Eugene, OR 97403

1 Dr. Alma E. Lantz
University of Denver
Denver Research Institute
Industrial Economics Division
Denver, CO 80210

1 Mr. Brian McNally
Educational Testing Service
Princeton, NJ 08540

1 Mr. A. J. Pesch, President
Eclectech Associates, Inc.
P. O. Box 178
North Stonington, CT 06359

1 Dr. Steven M. Pine
University of Minnesota
Department of Psychology
Minneapolis, MN 55455

1 Mr. Dennis J. Sullivan
c/o HAISC, Building 119, M.S. 2
P. O. Box 90515
Los Angeles, CA 90009

1 Dr. Patrick Suppes
Stanford University
Institute for Mathematical Studies
in the Social Sciences
Stanford, CA 94305

1 Dr. Carl R. Vest
Battelle Memorial Institute
Washington Operations
2030 M Street, NW
Washington, DC 20036

1 Dr. John J. Collins
Vice President
Essex Corporation
6305 Caminito Estrellado
San Diego, CA 92120

1 Mr. Charles R. Rupp
Advanced W/C Development Engineering
General Electric Company
100 Plastics Avenue
Pittsfield, MA 01201